LES DICTIONNAIRES DU SAVOIR MODERNE

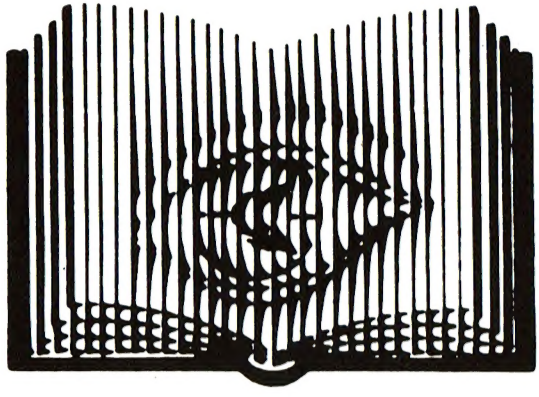
les mathématiques

modernes aussi bien que traditionnelles
expliquées dans un langage accessible à l'étudiant
comme à l'adulte qui se recycle

par Christiane de Bary, René Boirel, Pierre Buisson, Danielle Fèvre, Nicole Gautier
Maurice Glaymann, Michelle Pascal, François Russo, André Warusfel

tout le monde en a maintenant besoin,
et tout le monde peut les comprendre

Centre
d'Étude et de Promotion
de la Lecture



LES MATHEMATIQUES

comment lire le présent volume

Bien entendu, nous ne prétendons pas vous apprendre à lire les lettres de notre alphabet et les mots de notre vocabulaire.
Mais la lecture, pour l'« honnête homme » du XX^{e} siècle, c'est beaucoup plus: c'est, entre autres, une technique d'information aussi efficace que possible – aussi efficace, mais également aussi personnelle.
Votre droit de *lecteur adulte* est de lire le sujet traité dans cet ouvrage, en fonction non pas des préoccupations des auteurs, mais de *vos préoccupations,* de vos besoins, de vos désirs.
La structure de cet ouvrage, dénommée « Algo-livre », a été conçue pour permettre cette lecture libre et totale – sans pour autant dénaturer la pensée des auteurs.
Vous pouvez utiliser ce livre en pratiquant soit une lecture continue d'information générale en suivant l'ordre donné en page 5; soit une lecture sélective de recherche.

Vous trouverez en pages 6 et 7 du présent ouvrage le mode d'utilisation détaillé de l'Algo-livre.

En voici les principaux points:
□ Les neuf articles essentiels sont repérés sur la tranche du volume par des onglets permettant de les séparer avec le pouce; la table de ces articles, classés suivant un ordre progressif d'exposition, figure en page 5.
□ Ces articles essentiels, ainsi que tous les autres articles (au nombre de quelques centaines), sont classés, dans le corps du volume, alphabétiquement.
Il suffit, pour les trouver, de feuilleter ce livre en repérant les trois lettres imprimées en haut et à droite de chaque page (de droite) qui correspondent aux trois premières lettres des mots définis dans cette page.
□ Chaque mot important, donnant lieu à un article, est repéré tout le long de l'ouvrage par deux types de signaux:
– des *« soulignés »:* tout *souligné* sous un mot indique que ce mot est le sujet d'un article; si vous voulez en savoir plus sur ce terme, il vous suffit de rechercher l'article correspondant, classé, rappelons-le, alphabétiquement;
– des *« index intégrés »,* indiquant à la suite de chaque mot-titre d'un article les numéros des autres pages de l'ouvrage où figure ce même mot.

En « jouant » avec les « soulignés » et les « index intégrés », vous pouvez lire la totalité des textes de cet ouvrage, mais dans l'ordre correspondant à vos curiosités, à vos besoins personnels.

LES MATHEMATIQUES

ENSEMBLES DE NOMBRES

$\mathbb{N}$	des entiers naturels ($\mathbb{N}^* = 0$ exclu)
$\mathbb{Z}$	des entiers relatifs (entiers positifs et négatifs)
$\mathbb{D}$	des nombres décimaux
$\mathbb{Q}$	des nombres rationnels ($\mathbb{Q}^* = 0$ exclu)
$\mathbb{R}$	des nombres réels ($\mathbb{R}^* = 0$ exclu)
$\mathbb{C}$	des nombres complexes

LOGIQUE

$\neg$	négation de la proposition
et *ou* $\wedge$	conjonction des propositions
ou *ou* $\vee$	disjonction des propositions
$\Longrightarrow$	implique
$\Longleftrightarrow$	équivalent à
$\forall$	quel que soit (pour tout)
$\exists$	il existe au moins un

ENSEMBLES

$\varnothing$	ensemble vide
$\in$	appartient à (ou est élément de) ; $\notin$ n'appartient pas à
$\subset$	est inclus dans ; $\not\subset$ n'est pas inclus dans
$\cup$	réunion ; $\cap$ intersection
$\bigcup_{i \in I} E_i$	réunion de la famille $(E_i)_{i \in I}$, pour i appartenant à I
$\bigcap_{i \in I} E_i$	intersection de la famille $(E_i)_{i \in I}$
$E \times F$	produit cartésien de deux ensembles E et F
$\prod_{i \in I} E_i$	produit cartésien de la famille $(E_i)_{i \in I}$
E^n	produit cartésien de n ensembles égaux à E
$E \Delta F$	différence symétrique des deux ensembles E et F
$\complement_E F$	complémentaire de F dans E
$\aleph$	aleph (voir ce mot au dictionnaire)

RELATIONS

$x \equiv y$ (mod. $\mathcal{R}$)	x congru à y modulo $\mathcal{R}$
$E/\mathcal{R}$	ensemble quotient de l'ensemble E par la relation d'équivalence $\mathcal{R}$
$\leqslant$	inférieur ou égal à ; $<$ strictement inférieur à
$\geqslant$	supérieur ou égal à; $>$ strictement supérieur à
$x \mathcal{R} y$	x est lié à y par la relation $\mathcal{R}$
$f : x \longmapsto f(x)$	x a pour image $f(x)$ par l'application f
$f : E \longrightarrow F$	f est une application de l'ensemble E dans l'ensemble F

LOIS DE COMPOSITION

$\divideontimes$, T, $\perp$, $\circ$	étoile, truc, antitruc, rond

AUTRES SYMBOLES

$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$	somme des n éléments $x_1, x_2, \ldots, x_n$
$\prod_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n$	produit des n éléments $x_1, x_2, \ldots, x_n$
$\frac{df}{dx}$ *ou* $f'(x)$	dérivée de f
df	différentielle de f ; df_{x_0} *ou* $d_{x_0} f$ différentielle de f au point x_0
$\frac{\partial f}{\partial x_i}$ *ou* f'_{x_i}	i-ième dérivée partielle de f
$\int f(x)\, dx$	intégrale de f ; $\iint$ intégrale double ; $\iiint$ intégrale triple
$\lvert x \rvert$	valeur absolue de x ou module de x
$\lVert x \rVert$	norme de x
$n!$	factorielle n (lire le mot au dictionnaire)
$\overline{A}$	événement contraire de A, adhérence de la partie A d'un espace topologique
B	base d'un espace vectoriel
e	élément neutre, base de l'exponentielle
f, g, h	application, fonction
f'	dérivée de f ; f'' dérivée seconde ; $f^{(n)}$ dérivée n-ième
f^{-1}	application réciproque de l'application f
$f(x)$	image de l'élément x par l'application f
$f^{-1}(A)$	image réciproque de A par f
I	ensemble d'indices
I_E	application identique de E
K	corps, compact
M	matrice, module ; $\overline{M}$ conjuguée de la matrice M
$\vec{0}$	vecteur nul
P	matrice de passage, polynôme, probabilité
Q	forme quadratique
V, W	voisinage
$\mathcal{F}$	filtre
$\alpha, \beta, \lambda, \mu$	scalaires
σ	permutation, écart type
χ_A	(*khi* A) application caractéristique de la partie A
Ω	ensemble d'opérateurs

2

Nouvelle édition illustrée en couleurs

© Centre d'Étude et de Promotion de la Lecture
Paris, 1973 et 1975.

LES MATHEMATIQUES

LES DICTIONNAIRES DU SAVOIR MODERNE

Neuf articles essentiels présentés page ci-contre ▸

... et 500 termes classés par ordre alphabétique constituent ce dictionnaire de mathématiques

Comment se servir de cette encyclopédie et utiliser le système des onglets : voir pages suivantes

1 Chronologie des grandes évolutions des mathématiques

Historique
par François Russo

Historique

2 La théorie générale à la base des disciplines mathématiques; le fondement des mathématiques modernes

Les ensembles
par Pierre Buisson

Ensembles

3 Nombre entier, nombre relatif, nombre premier, nombre rationnel, nombre irrationnel, nombre transcendant, nombre complexe...

Les nombres
par André Warusfel

Nombres

4 Les lois de composition des ensembles. Monoïde, groupe, anneau, corps...

L'algèbre
par Maurice Glaymann

Algèbre

5 L'univers des fonctions. Intervalles, suites, continuité, limite, dérivées, intégrales...

L'analyse
par Christiane de Bary

Analyse

6 Le lien entre l'algèbre et la géométrie : espace vectoriel, applications linéaires, calcul matriciel...

L'algèbre linéaire
par Danielle Fèvre

Algèbre linéaire

7 La statistique descriptive; population, écart type, courbe de Gauss, coefficients de corrélation...
La statistique mathématique : les différentes lois et les sondages

Les statistiques
par Michelle Pascal

Statistiques

8 Les bases de la méthode statistique : événements, variable aléatoire, lois de probabilités, espérance mathématique, loi des grands nombres...

Les probabilités
par Michelle Pascal

Probabilités

9 De plus en plus importantes : des sciences exactes aux sciences humaines; aide à la décision, théorie des graphes, codage, modèles, structures...
Une nouvelle pédagogie

Les applications des mathématiques
par René Boirel

Applications

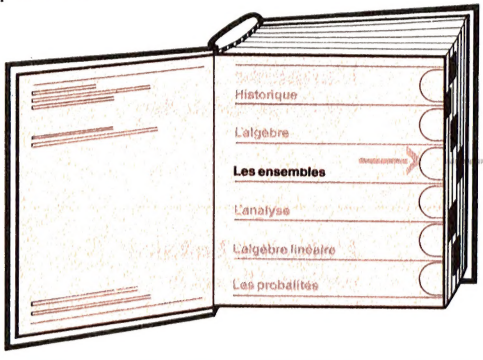
Comment utiliser cette

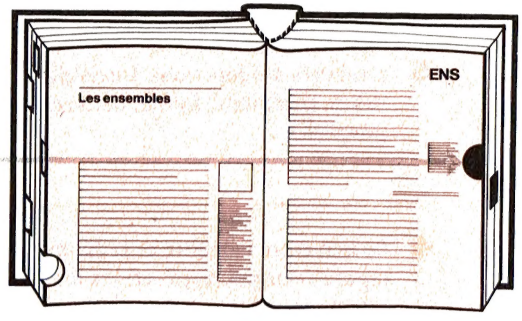
La structure originale de cet ouvrage, dénommée « Algo-livre », est conçue pour vous permettre de le consulter avec la même simplicité et le même profit, quels que soient votre objectif de lecture et votre niveau de connaissances.
Vous pouvez utiliser cette encyclopédie en pratiquant soit une lecture continue

1 Lecture continue d'information générale

Consultez la page précédente. Elle vous présente le sommaire d'un traité complet sur les mathématiques réparti en neuf chapitres : du chapitre premier « Historique » au chapitre neuvième « les Applications des mathématiques ». A chacune des neuf lignes de ce sommaire correspond sur la tranche verticale droite du volume une encoche qui permet de localiser la première page du chapitre correspondant.

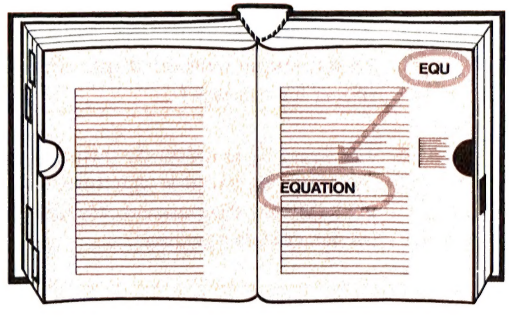
Par exemple, pour lire le chapitre sur « les Ensembles », il suffit, sur la page précédente, de déplacer un doigt horizontalement de la gauche vers la droite à hauteur de la ligne « les Ensembles » ; le doigt tombe dans une encoche qui permet automatiquement d'ouvrir le livre à la page 256, première page de l'article sur « les Ensembles ».





2 Lecture sélective de recherche

Pour vous informer sur un sujet particulier, utilisez le présent ouvrage à la façon d'un dictionnaire traditionnel. Tous les articles, aussi bien les études générales de plusieurs dizaines de pages (par exemple l'article sur « les Statistiques ») que les définitions de quelques lignes (par exemple « Équation », page 286), sont en effet classés alphabétiquement.
Il suffit, pour les trouver, de feuilleter ce livre en repérant les trois lettres imprimées en haut et à droite de chaque page (de droite), qui correspondent aux trois premières lettres des mots définis dans cette page.

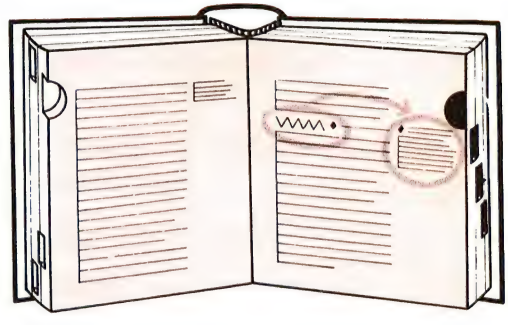


encyclopédie avec le meilleur profit

d'information générale, soit une lecture sélective de recherche; vous pouvez ou négliger ou retenir les renseignements spécialisés de nature technique ou scientifique; mais, quel que soit votre mode de lecture, il vous est toujours possible de repérer aisément les liaisons entre des notions apparemment distinctes, mais néanmoins interdépendantes.

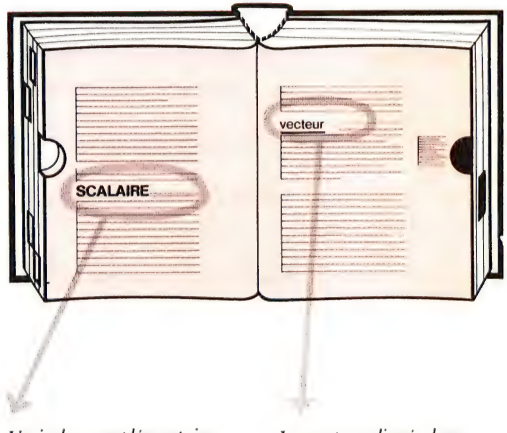
3 Textes annexes : notes marginales

Les articles de cet ouvrage ont été écrits sans jamais sacrifier à une vulgarisation déformante ou à des simplifications abusives, mais ils ont aussi été rédigés pour être compris, assimilés par tout lecteur doté d'une culture générale moyenne. Afin de faciliter votre tâche, nombre de renseignements, qui bien qu'utiles ne sont pas nécessairement indispensables, ont été séparés du texte principal, composés dans un caractère plus petit, placés en marge, à la hauteur du « mot signal ».
Ce « mot signal » est repéré par un losange noir.
On retrouve ainsi sous une forme moderne l'équivalent des « gloses » des manuscrits du Moyen Âge et des livres imprimés de la Renaissance.



4 Index général des mots

En rencontrant dans cet ouvrage certains termes vous éprouverez le besoin de consulter les autres pages du présent livre où ces mêmes termes sont cités : soit définis, soit développés, soit commentés.
Dans les ouvrages traditionnels, cette fonction est remplie par un index alphabétique placé en fin de volume.
La structure « Algo-livre » vous permet de trouver directement ces renseignements, sans l'obligation de compulser un cahier terminal d'index.
1. Les mots soulignés dans le cours du texte renvoient à ces mêmes mots, définis dans le dictionnaire.
2. Un index complémentaire est donné immédiatement après chacun des mots du dictionnaire.



Un index complémentaire est donné immédiatement après chacun des mots du dictionnaire.

Les mots soulignés dans le texte font l'objet d'une définition dans le dictionnaire.

Cette collection est publiée par
la Bibliothèque du Centre d'Étude
et de Promotion de la Lecture.
Conçue par François Richaudeau,
elle est réalisée sous la direction de Jean Feller,
assistés par Yvette Pesez.

Francine Achaz en assure le secrétariat de rédaction
et la fabrication
Ulrich Meyer, de Hollenstein Création,
les maquettes de reliure et de mise en pages.

La réalisation de cet ouvrage a été assurée par
Yvette Pesez

avec l'assistance scientifique de
Danielle Fèvre

Les auteurs sont :

Christiane de Bary — Agrégée de mathématiques, maître-assistant à la faculté de Metz

René Boirel — Docteur ès lettres, professeur à l'Institut national des sciences appliquées de Lyon

Pierre Buisson — Maître-assistant à l'université de Rouen

Danielle Fèvre — Agrégée de mathématiques, maître-assistant à la faculté de Metz

Maurice Glaymann — Maître-assistant à l'université Claude-Bernard de Lyon, directeur de l'Institut de recherche pour l'enseignement des mathématiques de Lyon

Michelle Pascal — Professeur, membre diplômé de l'Institut des actuaires français

François Russo — Polytechnicien, membre correspondant de l'Académie internationale d'histoire des sciences

André Warusfel — Professeur de mathématiques spéciales

et Nicole Gautier — Agrégée de mathématiques, professeur pour le dictionnaire

Gérard Guillerault ayant traité les mots « Topologie » et « Mathématiques du flou »

Les dessins ont été exécutés par
Evelyne Doucet,
Philippe Bully, Renée Espérance et Michel Galy.

Cet ouvrage étant un ouvrage collectif, chacun des auteurs n'assume que la responsabilité de son propre texte.

avertissement

Pour tous les ouvrages de cette collection, nous nous sommes assigné un double but : et satisfaire la curiosité des lecteurs non spécialistes — ce qui implique l'accessibilité des textes —, et faire le point des recherches modernes dans la discipline concernée.
Pour les mathématiques, nous avons opté pour une initiation à partir de bases élémentaires que l'on peut supposer connues de tous les lecteurs, avec une progression à la fois à travers le livre (indiquée en page 5) et en cours d'article.

Ce qui revient à conseiller au lecteur de lire :

1) Les chapitres principaux dans l'ordre donné page 5, *en allant des «Ensembles» aux «Applications des mathématiques» (pour chaque chapitre, les dernières pages sont parfois difficiles mais n'empêchent en aucun cas de suivre l'étape suivante). Ces grandes notions permettront aux parents de suivre et de comprendre l'enseignement moderne dispensé dans les cours secondaires ; aux curieux de s'intéresser à l'évolution des mathématiques ; et aux chercheurs, quels qu'ils soient, d'utiliser certaines connaissances dans leur propre discipline (et c'est la raison pour laquelle une grande importance a été donnée aux chapitres « Analyse », « Probabilités » et « Statistiques »).*

2) Les articles du dictionnaire proprement dit, *en tenant compte des nombreux renvois qui y sont notés. En effet, si un certain vocabulaire reste classique et familier, bon nombre de mots dépendent de notions modernes traitées dans les grands articles. On s'étonnera peut-être de ne pas trouver les noms des mathématiciens : il eût été finalement trop difficile d'expliquer leur œuvre sans charger le dictionnaire de notions souvent dépassées ; quelques notes biographiques ont été cependant données en références marginales.*
Nous signalons également que nous avons volontairement limité certains sujets — telles la géométrie moderne, la logique et la topologie —, qui sont des domaines de pointe des mathématiques et qui sont d'un autre niveau. Nous en avons néanmoins posé les principes pour la bonne règle de notre objectif : donner au lecteur le « savoir » du XX^e^ siècle.

Y. P.

A

ABAQUE : voir l'article « les Applications des mathématiques », pages 160, 162.

ABÉLIEN : voir Commutatif

ABSCISSE : voir Coordonnées

ABSORBANT ou PERMIS ◆ page 46.

◆ Il est conseillé de lire d'abord Composition (loi de).

Soit E un ensemble muni d'une loi de composition interne notée $\divideontimes$; un élément a de E est dit permis à droite (respectivement à gauche) si, pour tout élément x de E, on a :

$$a \divideontimes x = a \quad (\text{resp. } x \divideontimes a = a).$$

Un élément permis ou absorbant est un élément qui est à la fois permis à droite et permis à gauche.

Par exemple, dans l'ensemble des entiers relatifs $\mathbb{Z}$, muni de la multiplication, 0 est élément permis car, pour tout entier relatif n, on a :

$$0 \times n = n \times 0 = 0.$$

Dans l'ensemble des parties d'un ensemble E : $\mathcal{P}(E)$ muni de la loi intersection, l'ensemble vide $\varnothing$ est élément permis, car, pour toute partie X de E, on a :

$$X \cap \varnothing = \varnothing \cap X = \varnothing.$$

Dans un ensemble, il existe au plus un élément permis.

ABSURDE (Raisonnement par) pages 284, 328, 379.

Pour démontrer qu'une proposition A est fausse, on suppose qu'elle est vraie et on montre que cela conduit à une conclusion incompatible avec les hypothèses de départ.

Exemple : démontrons qu'il n'existe aucun nombre rationnel dont le carré soit égal à 2.

Supposons qu'il en existe un ; on le note $\frac{p}{q}$ (p et q étant deux nombres premiers entre eux) ; on a :

$$\frac{p^2}{q^2} = 2, \quad \text{soit} \quad p^2 = 2q^2$$

p^2 est donc un nombre pair, ce qui entraîne que p est pair ◆. D'où $p = 2k$ (k étant le nombre qui, multiplié par 2, donne p) ; alors :

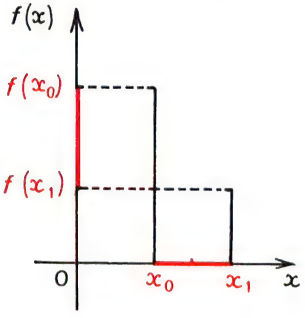
$$p^2 = 4k^2 = 2q^2, \text{ ce qui entraîne } q^2 = 2k^2,$$

ce qui prouve que q est pair aussi ; $q = 2k'$.

On aboutit à cette conclusion : p et q sont tous les deux multiples de 2, ce qui est contraire au fait que p et q soient premiers entre eux. L'hypothèse faite était donc fausse.

◆ En effet, le carré d'un nombre impair est encore un nombre impair.

ACCROISSEMENT pages 132, 328.



Si f est une fonction réelle de la variable réelle x et si x_0 et x_1 sont deux valeurs de cette variable telles que $f(x_0)$ et $f(x_1)$ existent, on appelle accroissement de la variable le nombre réel (positif ou négatif) $x_1 - x_0$ et accroissement correspondant de la fonction le nombre réel (positif ou négatif) $f(x_1) - f(x_0)$.

Sur le dessin, $x_1 - x_0$ est positif,
$f(x_1) - f(x_0)$ est négatif.

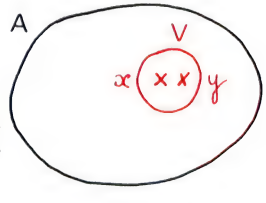
Théorème des accroissements finis : voir l'article « l'Analyse », page 132.

ACCUMULATION (Point d') ◆ page 14.

◆ Vocabulaire de topologie. Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire l'article « les Ensembles », Intervalle et Topologie.

Un point x d'une partie A d'un espace topologique E est un point d'accumulation de A si tout voisinage de x contient au moins un point de A autre que x.
N'importe quel point de l'intervalle [0, 1] est point d'accumulation, car tout intervalle entourant x contient des points de [0, 1] autres que x.

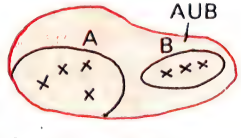
Si $\mathfrak{V}(x)$ est l'ensemble des voisinages de x, quel que soit le voisinage V appartenant à $\mathfrak{V}(x)$, il existe un élément y de A, différent de l'élément x, qui appartienne à l'intersection de V avec A :
$\forall\ V \in \mathfrak{V}(x), \exists y \in A \mid y \neq x, y \in V \cap A$.



◆ Lire aussi l'article « les Ensembles », page 265.

ADDITION, SOMME ◆ pages 17, 22, 24, 202, 368 à 389.

Si A et B sont deux ensembles disjoints dont les cardinaux sont respectivement m et n, par définition, le cardinal de la réunion A ∪ B est appelé la somme des deux cardinaux m et n et noté $m + n$ (m, n et $m + n$ sont trois entiers naturels). La loi de composition interne ainsi définie dans $\mathbb{N}$ est appelée addition. Cette loi commutative, associative, est telle que 0 soit élément neutre



Card A = m Card B = n
Card (A ∪ B) = $m + n$

De façon plus générale, on a l'habitude d'appeler addition et de noter + toute loi de composition interne dans un ensemble E associative et commutative ; le composé de deux éléments x et y est noté $x + y$ et appelé somme des deux éléments.

La loi étant associative, on peut définir successivement le composé de 3 éléments, ..., de n éléments que l'on note $x_1 + x_2 + \ldots + x_n$ et qu'on appelle somme de ces éléments.

ADHÉRENCE ◆

◆ Vocabulaire de topologie. Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

Si A est une partie d'un espace topologique E, l'intersection de tous les fermés de E contenant A est un fermé de E contenant A (c'est le plus petit de tous les fermés contenant A puisqu'il est contenu dans chacun d'entre eux) ; ce fermé est appelé adhérence de A et noté $\overline{A}$.

Exemple

L'adhérence d'un intervalle (a, b) de la droite numérique $\mathbb{R}$ est l'intervalle fermé $[a, b]$.

● Si A est un fermé de E, alors $A = \overline{A}$, et réciproquement.

● Si A et B sont deux parties de E, telles que A soit contenue dans B, alors l'adhérence de A est contenue dans l'adhérence de B :

$$A \subset B \Longrightarrow \overline{A} \subset \overline{B}.$$

● Si A et B sont deux parties de E, l'adhérence de la réunion $A \cup B$ est égale à la réunion des adhérences de A et de B :

$$\overline{A \cup B} = \overline{A} \cup \overline{B}.$$

Mais l'adhérence de l'intersection $A \cap B$ n'est pas égale à l'intersection des adhérences de A et de B ; cette intersection $\overline{A} \cap \overline{B}$ est seulement contenue dans l'adhérence de $A \cap B$:

$$\overline{A} \cap \overline{B} \subset \overline{A \cap B}.$$

Exemple

Soient A et B, respectivement, les intervalles ouverts $]a, b[$ et $]b, c[$ de la droite numérique $\mathbb{R}$.

L'adhérence de A : $\overline{A}$ est l'intervalle fermé $[a, b]$;

l'adhérence de B : $\overline{B}$ est l'intervalle fermé $[b, c]$.

Donc, l'intersection des adhérences est réduite au point b :

$$\overline{A} \cap \overline{B} = \{ b \}.$$

L'intersection $A \cap B$ est vide (aucun élément commun pour ces deux intervalles ouverts), donc l'adhérence de $A \cap B$ est vide ; aussi :

$$\overline{A \cap B} = \varnothing.$$

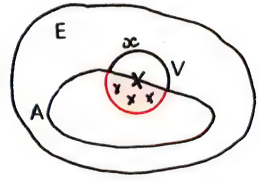
● L'adhérence d'une partie A de E est aussi l'ensemble des points adhérents à A.

ADHÉRENT ◆

◆ Vocabulaire de topologie. Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

Un point x d'un espace topologique E est dit adhérent au sous-ensemble A de E si l'intersection de A avec n'importe lequel des voisinages de x n'est pas vide. $\mathfrak{V}(a)$ étant l'ensemble des voisinages de a :

$$\forall\, V \in \mathfrak{V}(a),\ V \cap A \neq \varnothing.$$



Exemple

Dans $\mathbb{R}$ (ensemble des nombres réels), zéro est adhérent à l'ensemble A des inverses des nombres entiers naturels non nuls :

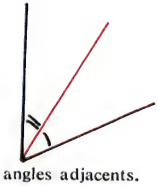
$$A = \left\{ \frac{1}{n} \;\middle|\; n \in \mathbb{N}^* \right\}$$

(un point isolé de A, un point d'accumulation de A sont adhérents à A).

ADJACENT

« Petit Larousse » : qui se touche, attenant, contigu.

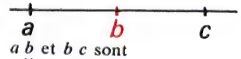
Angles adjacents : angles ayant même sommet, un côté commun et situés de part et d'autre de ce côté.



angles adjacents.

De façon générale ◆, on dit que deux parties d'un espace topologique sont adjacentes si leurs frontières ont au moins un point commun et si aucune des deux parties n'a de point à l'intérieur de l'autre. Un exemple simple est celui de segments adjacents ab et bc d'une droite : leurs frontières ont un seul point commun, le point b, qui est extrémité commune aux deux segments (ce point b peut ne pas appartenir aux segments si ce ne sont pas des intervalles fermés ◆).

◆ Lire d'abord Topologie.



$a\,b$ et $b\,c$ sont adjacents.

Dans le cas d'un ensemble totalement ordonné, deux sous-ensembles sont adjacents si la borne supérieure de l'un est la borne inférieure de l'autre (les deux sous-ensembles peuvent être disjoints ou non). Par exemple, dans $\mathbb{R}$ (ensemble des nombres réels), les deux sous-ensembles A et B définis de la façon suivante :

◆ Voir Intervalle.

A est l'ensemble de tous les réels compris entre 0 et 2 (strictement) :

$$A = \left\{ x \;\middle|\; x \in \mathbb{R},\ 0 < x < 2 \right\}$$

B est l'ensemble des réels compris strictement entre 2 et 3 :

$$B = \left\{ x \;\middle|\; x \in \mathbb{R},\ 2 < x < 3 \right\}$$

sont adjacents, car la borne supérieure de A qui est 2 est aussi la borne inférieure de B (A et B sont néanmoins disjoints).

(Dans ce cas précis, on retrouve exactement la définition de segments adjacents puisqu'on peut représenter $\mathbb{R}$ par une droite et tout intervalle de $\mathbb{R}$ par un segment de cette droite.)

On peut remarquer que la définition donnée, dans le cas de nombres réels, peut être remplacée par les conditions suivantes :

A et B sont adjacents si et seulement si :

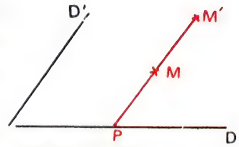
1) tout élément de A est inférieur à tout élément de B ;

2) quel que soit le nombre positif ε (donné à priori aussi petit que l'on veut), on peut toujours trouver au moins un élément y de B et un élément x de A tels que la différence $y - x$ soit plus petite que ε.

AFFINE : voir l'article « l'Algèbre linéaire », pages 62, 71 et la page 318.

AFFINITÉ (Géométrie plane)

Soient D et D′ deux droites non parallèles et k un nombre réel non nul. On appelle affinité d'axe D, de directrice D′ et de rapport k l'application qui, à chaque point M du plan, fait correspondre le point M′ tel que le vecteur $\overrightarrow{PM'}$ soit égal à k fois le vecteur $\overrightarrow{PM}$: $\overrightarrow{PM'} = k.\overrightarrow{PM}$
(P étant le point d'intersection de D avec la parallèle à D′ menée par M : voir figure ci-contre).
Si les droites D et D′ sont perpendiculaires, on dit que l'affinité est orthogonale.



Affinité d'axe D, de directrice D′ et de rapport 2.

AIRE ◆ page 317.

◆ La notion d'aire (comme celle de volume) est un cas particulier de la notion de mesure.

Si on a choisi une unité de longueur et si la longueur des côtés d'un carré a pour mesure a (avec cette unité), l'aire de la portion de plan limitée par ce carré est le nombre ◆ positif égal à $a \times a = a^2$. L'aire de la portion de plan limitée par un rectangle dont les côtés (mesurés avec la même unité) ont pour longueurs respectives a et b est le nombre positif $a \times b$. L'aire d'un triangle de côté a et de hauteur h est égale à $\frac{a \times h}{2}$.

◆ Ne pas confondre aire et surface : l'aire est un nombre réel positif, et une surface est un ensemble de points.

L'aire d'un polygone est obtenue en décomposant ce polygone en triangles et en faisant la somme des aires de ces triangles.
L'aire d'une portion du plan limitée par une courbe fermée (C) est la limite de l'aire d'un polygone inscrit dans cette courbe lorsque le nombre de ses côtés augmente indéfiniment.
On peut calculer l'aire d'un tel domaine à l'aide d'une intégrale simple ou double ◆.
L'aire d'un domaine plan est, quand elle existe, un nombre positif ou nul.

◆ Voir l'article « l'Analyse », page 134.

— L'aire de l'ensemble vide $\varnothing$ est nulle.
— $A \subset E \Longrightarrow$ aire de $A \leqslant$ aire de E.
— $A \cap B = \varnothing$; aire de $(A \cup B)$ = aire de A + aire de B.

AJUSTEMENT : voir les pages 486, 491, 493, 494, 502 à 507.

ALEPH

Lettre hébraïque qui, munie d'un indice, sert à désigner les cardinaux des ensembles infinis (nombres transfinis) : aleph zéro $\aleph_0$ est le cardinal de $\mathbb{N}$ (ensemble des entiers naturels), $\aleph_1$ est celui de $\mathbb{R}$ (ensemble des nombres réels).

ALGÈBRE : voir l'article page suivante et les pages 194, 526.

Partie des mathématiques (généralisation de l'arithmétique) dont l'objet était, au départ, l'étude des opérations sur les nombres entiers, rationnels, réels ou complexes ; cet objet devint par la suite l'étude de lois de composition (entre éléments qui n'étaient plus nécessairement des nombres) et de structures (groupe, anneau, corps...).

ALGÈBRE LINÉAIRE : voir l'article page 55.

L'algèbre

par Maurice Glaymann

Les premiers mathématiciens exercèrent leur talent et leur imagination à découvrir des *propriétés particulières* des nombres et des figures géométriques ; ils n'étudiaient jamais des ensembles de nombres et encore moins de lois de composition sur de tels ensembles. Un premier pas vers l'abstraction et, ce faisant, un progrès considérable ont été accomplis lorsque les mathématiciens ont pris la liberté de remplacer des nombres par des lettres ou des symboles et qu'ils ont appris à calculer sur ces objets : l'algèbre venait de naître... Dès lors, elle a permis de se libérer de l'étude de problèmes particuliers : le mathématicien disposait d'un outil si efficace qu'il pouvait résoudre d'un seul coup toute une famille de problèmes. On fait abstraction des nombres, on ne raisonne plus que sur des symboles par le jeu de relations. Cette idée s'est alors généralisée : les symboles ne représentent plus des nombres mais des objets quelconques et souvent non précisés. Dans cet article, nous allons apprendre à *composer :* en prenant un ensemble E dont les éléments peuvent être « n'importe quoi », nous définirons une ou plusieurs lois et nous les étudierons. La notion de loi de composition, bien que féconde, ne suffit pas, à elle seule, pour découvrir des résultats généraux. Il faut dépasser ce stade. On se donne un ensemble E, des lois de composition sur E et on *impose* certaines conditions à ces lois : *commutativité*, *associativité*, *existence d'un élément neutre*, *existence d'éléments symétriques*, *distributivité*, etc. Nous obtenons ainsi une *structure*, notion qui permettra une véritable économie de pensée : en effet, un résultat vrai dans une structure sera aussi vrai dans tout modèle d'une telle structure, sans qu'il soit besoin pour autant de le redémontrer pour chaque modèle...

Une structure particulièrement belle et féconde est celle de *groupe* ; elle a permis de fonder de nombreuses théories modernes et elle trouve des applications dans un grand nombre de domaines, en particulier la physique théorique. C'est Évariste Galois ◆ qui, en 1831, a montré l'intérêt de ce concept et a ainsi ouvert la voie aux plus grandes idées mathématiques contemporaines.

Maurice Glaymann.
Né en 1931.
Maître-assistant à l'université Claude-Bernard de Lyon, il consacre l'essentiel de ses activités à l'enseignement des mathématiques.
Président de l'Association des professeurs de mathématiques de 1966 à 1968, il est actuellement directeur de l'Institut de recherche pour l'enseignement des mathématiques de Lyon.

◆ E. Galois (1812-1832) : mathématicien français dont le génie précoce a été méconnu de son vivant. Pionnier de la notion de groupe de l'équation, il a également introduit l'idée de structure.

Loin de restreindre notre liberté, les structures — en mathématique — fournissent des méthodes générales de recherches, elles élargissent notre champ d'action et excitent notre imaginagion.

LOI DE COMPOSITION

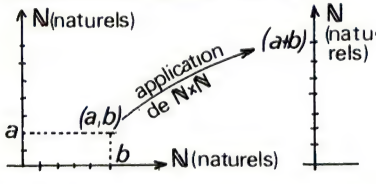
Partons de l'ensemble $\mathbb{N}$ des nombres *naturels*. La notion d'addition dans $\mathbb{N}$ est certainement aussi ancienne que celle de nombre. Que faisons-nous lorsque nous effectuons une addition ? Nous choisissons deux éléments de $\mathbb{N}$, 2 et 5 par exemple, et, à ces deux éléments, nous associons un troisième naturel, appelé somme de ces deux naturels : 7, et nous écrivons :

$$2 + 5 = 7.$$

Ainsi, au couple (2, 5) de $\mathbb{N} \times \mathbb{N}$ (lire $\mathbb{N}$ croix $\mathbb{N}$) ◆, nous associons l'élément 7 de l'ensemble $\mathbb{N}$.

◆ Voir la définition du couple dans l'article « les Ensembles », page 268.

Application de N×N vers N



Nous sommes en présence d'une application de $\mathbb{N} \times \mathbb{N}$ vers $\mathbb{N}$; une telle application est appelée *loi de composition* dans $\mathbb{N}$.

La loi de composition est ici l'addition dans $\mathbb{N}$.
Plus généralement, on appelle loi de composition sur un ensemble E une *application f* du produit cartésien $E \times E$ (E croix E) vers E, et on la note :

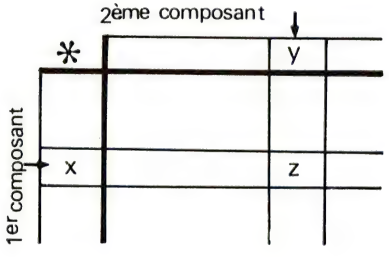
$$f : E \times E \longrightarrow E$$
$$f : (x, y) \longrightarrow z$$

z est l'image de (x, y) par l'application f.
En général, on introduit un symbole pour désigner la loi de composition, une *étoile* par exemple :

$$z = x \mathbin{\divideontimes} y.$$

Le symbole $\divideontimes$ est le signe opératoire de la loi de composition : pour les lois de composition classiques, comme l'addition et la multiplication, nous utiliserons les symboles $+$ et $\times$.
Dans l'écriture $z = x \divideontimes y$, z est le *composé* de x et de y ; x est le premier composant et y le second composant.
Attirons l'attention sur le fait que, dans la définition d'une loi de composition, il est question d'une application.

Ainsi, l'addition dans $\mathbb{N}$ est une loi de composition, car à tout couple de naturels (a, b) de $\mathbb{N} \times \mathbb{N}$, nous pouvons faire correspondre l'élément $a + b$ qui en est la somme ; par contre, la soustraction dans $\mathbb{N}$ n'est pas une loi de composition, car au couple (a, b) nous ne pouvons associer l'élément $a - b$ que si $a \geqslant b$. On dit que la soustraction est une *opération* dans $\mathbb{N}$.
Du point de vue pratique, le problème est de déterminer z ; en général, la définition même de la loi de composition donne un procédé pour calculer z. Dans certains cas, nous pouvons donner une table des composés (cela n'est possible que si l'ensemble E est fini).



On dispose, dans la première colonne de gauche et dans la première ligne, les éléments de E(x et y) et on écrit le composé z de x et de y à l'intersection de la ligne contenant x et de la colonne contenant y. Une telle table s'appelle la *table de Pythagore* de la loi de composition ✳.

Exemple

Voici la table de Pythagore d'une loi de composition notée ✳ dans l'ensemble $E = \{ a, b, c, d \}$. Cette table, composée de façon tout à fait arbitraire, définit complètement la loi de composition.

✳	a	b	c	d
a	a	c	b	d
b	b	d	a	c
c	d	b	c	a
d	c	a	d	b

Cherchons quel est le composé de b par c. Il se trouve à l'intersection de la ligne b (1er composant) et de la colonne c (2e composant) ; c'est a :

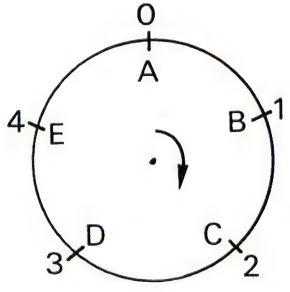
$(b \text{ ✳ } c) = a$;

de même, $(c \text{ ✳ } a) = d$, etc.

Si l'on veut effectuer des calculs, il est bon d'avoir la table sous les yeux ou de l'apprendre par cœur ; ce n'est pas un but en soi : *l'étude des propriétés de la loi de composition est beaucoup plus importante.*
Donnons quelques exemples de lois de composition.

Exemple a)

Voici un exemple particulier que nous allons développer et que nous exploiterons dans la suite. Nous donnerons ici une présentation très simple d'un concept qui joue un rôle important en algèbre.



Prenons un disque en carton qui porte les lettres A, B, C, D, E régulièrement espacées ; le disque peut tourner autour de son centre. La lettre A est devant le chiffre 0. Faisons tourner le disque d'un cinquième de tour. La lettre A vient se placer devant le chiffre 1. Nous aurions obtenu le même résultat en faisant tourner le disque d'un tour et d'un cinquième de tour, de deux tours et d'un cinquième de tour, etc.

Désignons toutes ces rotations par le symbole 1.
De même, le symbole 2 désigne les rotations de deux cinquièmes de tour, de un tour et de deux cinquièmes de tour, etc. On définit de la même façon 3 et 4.
Les rotations de un tour complet, deux tours, etc., sont désignées par 0, puisque l'on revient au point de départ. Définissons alors une loi de composition sur l'ensemble de ces rotations :

$$C = \{ 0, 1, 2, 3, 4 \}$$

Si nous faisons subir au disque une première rotation $r_1 = 2$ (deux cinquièmes de tour), puis une seconde rotation $r_2 = 1$ (un cinquième de tour), la position du disque après cette seconde rotation peut être atteinte à partir de la position initiale par la rotation $r_3 = 3$ (trois cinquièmes de tour).
r_3 est la composée de r_1 et de r_2 ; nous écrirons :

$$r_3 = r_1 \ast r_2 \quad \text{ou} \quad 3 = 1 \ast 2.$$

De même, $2 \ast 3 = 0$ (voir figure).
Nous sommes en présence d'une loi de composition sur C, dont voici la table de Pythagore:

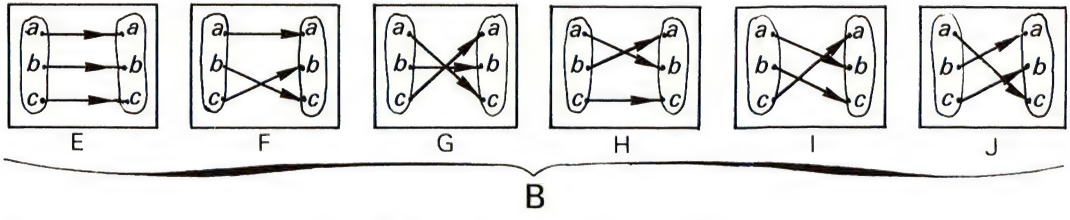
$\ast$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Pour obtenir cette table, il suffit d'effectuer avec le disque toutes les rotations associées deux à deux, soit 25 manipulations.

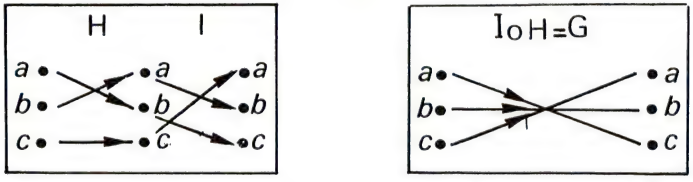
En prenant maintenant quatre lettres A, B, C, D régulièrement espacées, vous définirez quatre rotations 0, 1, 2 et 3 et vous pourrez construire une table analogue à la précédente.

Exemple b)

Voici un ensemble $\{ a, b, c \}$.
Il existe six *bijections* de cet ensemble sur lui-même :



Désignons par B l'ensemble $\{ E, F, G, H, I, J \}$ de ces bijections ; à deux éléments de B, nous pouvons en associer un troisième par la loi de composition des applications. La composée de H et de I, par exemple, notée I ○ H, est l'application *unique* qui a le même effet que l'application H suivie de l'application I.



Appliquer la bijection H puis la bijection I à l'ensemble $\{ a, b, c \}$ revient à faire une seule bijection, comme le montre le tableau :

$$\begin{array}{lll} \overset{H}{a} \longmapsto b \overset{I}{\longmapsto} c & & \overset{I \circ H}{a} \longmapsto c \\ b \longmapsto a \longmapsto b & \text{ou encore} & b \longmapsto b \\ c \longmapsto c \longmapsto a & & c \longmapsto a \end{array}$$

Cette bijection coïncide avec la bijection de G ; nous noterons : $G = H \circ I$.
Déterminons de même H ○ I :

$$\begin{array}{lll} \overset{I}{a} \longmapsto b \overset{H}{\longmapsto} a & & \overset{I \circ H}{a} \longmapsto a \\ b \longmapsto c \longmapsto c & \text{ou encore} & b \longmapsto c \\ c \longmapsto a \longmapsto b & & c \longmapsto b \end{array}$$

H ○ I est donc la bijection F.
Notez au passage que H ○ I et I ○ H sont deux bijections différentes.

○	E	F	G	H	I	J
E	E	F	G	H	I	J
F	F	E	J	I	H	G
G	G	I	E	J	F	H
H	H	J	I	E	G	F
I	I	G	H	F	J	E
J	J	H	F	G	E	I

Par des calculs analogues, nous obtenons la table de Pythagore ci-contre.

L'examen attentif des deux tables que nous venons de construire met en évidence certaines particularités de ces deux lois de composition ; l'intérêt d'une loi de composition réside dans ses propriétés. Nous allons étudier certaines propriétés.

COMMUTATIVITÉ

Dans l'exemple *a)*, vous constatez que, quels que soient les éléments x et y de l'ensemble C,

$$x \divideontimes y = y \divideontimes x$$

Par exemple, $3 \divideontimes 0 = 0 \divideontimes 3$.

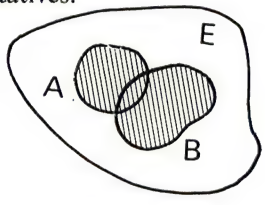
✻	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

La table de Pythagore de la loi de composition est *symétrique* par rapport à sa *diagonale principale*.
La loi de composition ✻ dans l'ensemble C est *commutative*.

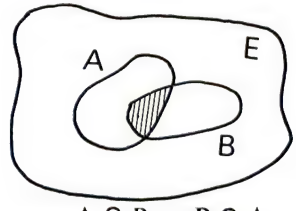
Plus généralement, « la loi de composition notée $\perp$, dans un ensemble E, est commutative » signifie, quel que soit le couple (x, y) de $E \times E$: $x \perp y = y \perp x$.
Si, pour une loi de composition notée T, dans un ensemble F, il existe au moins un couple (x, y) de $F \times F$ tel que $x \mathrm{T} y \neq y \mathrm{T} x$, la loi de composition n'est pas commutative. C'est le cas dans l'exemple *b)* ; nous avons vu que $H \circ I \neq I \circ H$.

Exemples et contre-exemples

1) L'addition et la multiplication dans $\mathbb{N}$ sont commutatives.
2) Si $\mathcal{P}(E)$ est l'ensemble des parties d'un ensemble E, la *réunion* et l'*intersection* sont des lois de composition sur $\mathcal{P}(E)$; elles sont commutatives.



$A \cup B = B \cup A$



$A \cap B = B \cap A$

3) L'*exponentiation* sur $\mathbb{N}$ n'est pas commutative. En général :

$$a^b \neq b^a.$$

Ainsi, par exemple : $2^3 = 8$ et $3^2 = 9$.

ÉLÉMENT NEUTRE

Reprenons l'exemple *a)* ; l'élément 0 joue un rôle très particulier :

$$\begin{aligned} 0 \ast 0 &= 0 \\ 0 \ast 1 &= 1 \ast 0 = 1 \\ 0 \ast 2 &= 2 \ast 0 = 2 \\ 0 \ast 3 &= 3 \ast 0 = 3 \\ 0 \ast 4 &= 4 \ast 0 = 4 \end{aligned}$$

Autrement dit, quel que soit l'élément x de C :

$$0 \ast x = x \quad \text{et} \quad x \ast 0 = x$$

De même, dans l'exemple *b)*, l'élément E joue un rôle analogue :
Quel que soit l'élément x de B :

$$x \circ E = x \quad \text{et} \quad E \circ x = x$$

Plus généralement, soient un ensemble A, une loi de composition notée $\times$ définie sur A et un élément e de E . « e est *élément neutre* pour la loi de composition $\times$ » signifie :
Quel que soit l'élément x de E :

$$x \ast e = x \quad \text{et} \quad e \ast x = x$$

Exemples

1) 0 est l'élément neutre pour l'addition sur $\mathbb{N}$.
2) 1 est l'élément neutre pour la multiplication sur $\mathbb{N}$.

3) Sur l'ensemble $\mathfrak{P}(E)$ des parties de l'ensemble E, E est élément neutre pour l'intersection : quelle que soit la partie A de E,

$$A \cap E = A \quad \text{et} \quad E \cap A = A$$

De même, la partie vide $\varnothing$ est élément neutre pour la réunion : quelle que soit la partie A de E,

$$A \cup \varnothing = A \quad \text{et} \quad \varnothing \cup A = A$$

Il ne faut pas croire que pour toute loi de composition il existe un élément neutre. Le lecteur vérifiera que pour la loi de composition $\divideontimes$ sur $\{ a, b, c, d \}$ de la page 18, aucun élément n'est élément neutre.
Cependant, lorsqu'il existe un élément neutre, il est *unique*.
Supposons que pour la loi de composition notée $\divideontimes$, définie sur l'ensemble E, existent *deux* éléments neutres e et u.
Désignons par f le composé de e et de u :

$$f = e \divideontimes u$$

Comme e est élément neutre : $f = e \divideontimes u = u$.

Comme u est élément neutre : $f = e \divideontimes u = e$.

Il en résulte que : $e = u$.

Ainsi, pour une loi de composition donnée, l'élément neutre lorsqu'il existe est unique.

ÉLÉMENTS SYMÉTRIQUES

Revenons à l'exemple *a)*. Nous venons de voir que 0 est l'élément neutre de la loi de composition $\divideontimes$ sur $\{ 0, 1, 2, 3, 4 \}$.
Vous constatez sur la table de Pythagore :

$$0 \divideontimes 0 = 0$$
$$1 \divideontimes 4 = 4 \divideontimes 1 = 0$$
$$2 \divideontimes 3 = 3 \divideontimes 2 = 0$$

Deux cinquièmes de tour et trois cinquièmes de tour sont des éléments symétriques puisque leur composée 2×3 est la rotation nulle. Autrement dit, quel que soit l'élément x, il existe un élément y tel que

$$x \divideontimes y = 0 \quad \text{et} \quad y \divideontimes x = 0$$

Plus généralement, soit un ensemble A et soit une loi de composition notée $\divideontimes$ définie sur A ; e est l'élément neutre pour cette loi de composition $\divideontimes$. Les éléments x et y de A sont dits des *éléments symétriques* pour la loi de composition $\divideontimes$, si :

$$x \divideontimes y = e \quad \text{et} \quad y \divideontimes x = e$$

On ne peut parler d'éléments symétriques que s'il existe un élément neutre pour la loi de composition considérée.
Inversement, même si l'élément neutre existe, il peut arriver que certains éléments ne possèdent pas de symétrique. C'est le cas pour l'ensemble des naturels $\mathbb{N}$ muni de l'addition ; seul 0 possède un symétrique : lui-même. Les autres éléments n'ont pas de symétrique. C'est pour cette raison que l'on est conduit à construire (par *symétrisation*) l'ensemble $\mathbb{Z}$ des entiers, dans lequel tout élément possède un symétrique pour l'addition.
Un élément peut être son propre symétrique. Ainsi, dans l'exemple *b)*, vous avez $H \circ H = E$, $G \circ G = E$, $F \circ F = E$.

ASSOCIATIVITÉ

Considérons l'addition sur $\mathbb{N}$. Lorsque nous écrivons $(7 + 5) + 2$, il est convenu que l'on commence par effectuer l'opération entre parenthèses : $7 + 5 = 12$.
La seconde opération consiste alors à effectuer $12 + 2$.

Ainsi, $(7 + 5) + 2 = 12 + 2 = 14$.

De même, l'écriture $7 + (5 + 2)$

conduit à calculer $5 + 2$, puis $7 + 7$.
On constate alors que

$$(7 + 5) + 2 = 7 + (5 + 2)$$

On démontre que, quels que soient les naturels a, b, c :

$$(a + b) + c = a + (b + c)$$

L'addition dans $\mathbb{N}$ est *associative*.
Plus généralement, soit un ensemble E muni de la loi de composition notée $\divideontimes$. « La loi de composition $\divideontimes$ est *associative* » signifie :
Quels que soient les éléments x, y, z de E :

$$(x \divideontimes y) \divideontimes z = x \divideontimes (y \divideontimes z)$$

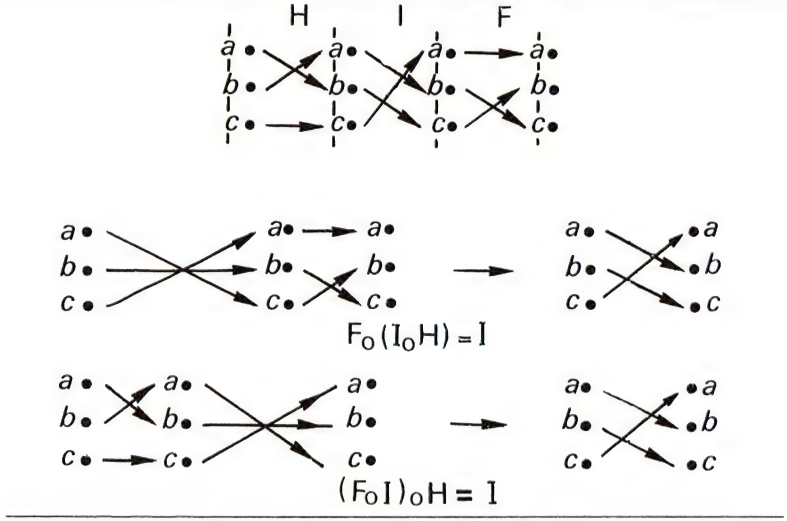
Lorsqu'une loi de composition $\divideontimes$ sur un ensemble E est associative, il est inutile de conserver des parenthèses pour écrire le composé de trois éléments (ou plus) : l'élément

$$(x \divideontimes y) \divideontimes z = x \divideontimes (y \divideontimes z)$$

se note $x \divideontimes y \divideontimes z$.

Exemples

1) La composition des bijections sur un ensemble E est une loi de composition associative. Reprenons l'exemple *b)* :



Nous avons bien $F \circ (I \circ H) = (F \circ I) \circ H$.
2) La multiplication sur $\mathbb{N}$ est associative.
3) L'intersection et la réunion sur $\mathfrak{P}(E)$ des parties de l'ensemble E sont associatives.

Contre-exemple : pour démontrer qu'une loi de composition notée $\divideontimes$ n'est pas associative, il suffit de trouver trois éléments a, b, c tels que

$$(a \divideontimes b) \divideontimes c \neq a \divideontimes (b \divideontimes c).$$

Ainsi, l'exponentiation dans $\mathbb{N}$ n'est pas associative, car

$$(2^3)^2 = 64 \quad \text{et} \quad 2^{(3^2)} = 512.$$

ÉLÉMENT RÉGULIER

E est un ensemble muni d'une loi de composition notée $\divideontimes$; a, x et y sont trois éléments de E.
Si, quels que soient les éléments x et y, l'égalité

$$a \divideontimes x = a \divideontimes y \quad \text{entraîne} \quad x = y,$$

alors l'élément a est *régulier à gauche* pour la loi de composition $\divideontimes$.
De même, si, quels que soient les éléments x et y, l'égalité

$$x \divideontimes a = y \divideontimes a \quad \text{entraîne} \quad x = y,$$

alors l'élément a est *régulier* à *droite* pour la loi de composition $\divideontimes$.
Si un élément a est régulier à gauche et à droite, on dit qu'il est *régulier* pour la loi de composition $\divideontimes$.

En particulier, si la loi de composition $\divideontimes$ est associative, si elle admet un élément neutre et si l'élément a admet un symétrique a', alors il est régulier. En effet :

$$a \divideontimes x = a \divideontimes y$$

entraîne $a' \divideontimes (a \divideontimes x) = a' \divideontimes (a \divideontimes y)$
compte tenu de l'associativité

$$(a' \divideontimes a) \divideontimes x = (a' \divideontimes a) \divideontimes y$$

a' étant le symétrique de a, $a' \divideontimes a = e$, élément neutre.

Donc $$e \divideontimes x = e \divideontimes y$$

d'où $$x = y.$$

On démontre de même que

$$x \divideontimes a = y \divideontimes a \quad \text{entraîne} \quad x = y.$$

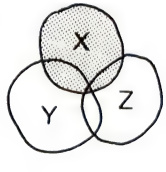
Donc a est régulier.

Contre-exemple. 0 n'est pas régulier pour la multiplication des entiers :

$$0 \times 12 = 0 \times 13 \quad \text{n'entraîne pas} \quad 12 = 13.$$

DISTRIBUTIVITÉ

Jusqu'à présent, nous nous sommes limités à un ensemble muni d'une seule loi de composition. Il est naturel de munir un ensemble de plusieurs lois de composition. Ainsi, par exemple, pour l'ensemble des naturels $\mathbb{N}$, vous connaissez deux lois de composition : l'addition et la multiplication.
De même, sur $\mathfrak{P}(E)$ ensemble des parties d'un ensemble E, vous connaissez deux lois de composition : l'intersection et la réunion. X, Y et Z étant trois parties de E, nous avons :



$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$$

la réunion est *distributive* sur l'intersection.

De même, $$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$$

l'intersection est distributive sur la réunion.
Dans l'ensemble $\mathbb{N}$, a, b et c étant trois naturels,

$$a \times (b + c) = (a \times b) + (a \times c)$$

la multiplication est distributive sur l'addition.
Par contre,

$$a + (b \times c) \neq (a + b) \times (a + c)$$

l'addition n'est pas distributive sur la multiplication.

Plus généralement, E est un ensemble muni de deux lois de composition notées $\perp$ et T. « La loi de composition $\perp$ est *distributive à gauche* sur la loi de composition T » signifie :
Quels que soient les éléments a, b, c de E

$$a \perp (b \mathrm{T} c) = (a \perp b) \mathrm{T} (a \perp c).$$

« La loi de composition $\perp$ est *distributive à droite* sur la loi de composition T » signifie :
Quels que soient les éléments a, b, c de E

$$(b \mathrm{T} c) \perp a = (b \perp a) \mathrm{T} (c \perp a).$$

Si la loi de composition $\perp$ est distributive à gauche et à droite sur la loi de composition T, on dit qu'elle est *distributive* sur la loi de composition T.

LES STRUCTURES

C'est en 1831 que le mathématicien Evariste Galois montre l'importance de la notion de structure de groupe en algèbre, pour étudier la résolution des équations. Quarante ans plus tard, Felix Klein ◆ montre à son tour, dans le « Programme d'Erlangen », tout le parti que l'on peut tirer de ce concept en géométrie. Depuis, la théorie des groupes n'a fait que s'enrichir et nous pouvons ajouter en outre qu'elle trouve tout naturellement sa place aujourd'hui dans les programmes scolaires ; l'expérience montre que, bien préparées, ces notions sont à la portée d'un enfant de treize ans !
Nous étudierons ici les structures les plus importantes : monoïde, groupe, anneau et corps.

◆ F. Klein (1849-1925) : mathématicien allemand auquel on doit d'importants travaux sur l'équation hypergéométrique, les fonctions abéliennes, l'application de la théorie des groupes à la géométrie et les fonctions elliptiques.

> Dès que l'on a défini une ou plusieurs lois de composition dans un ensemble, on dit que cet ensemble possède une *structure*.
> La structure est caractérisée par les propriétés que l'on impose aux lois de composition.

MONOIDE

E est un ensemble muni d'une loi de composition notée $\divideontimes$. « Le couple (E, $\divideontimes$) est un *monoïde* » signifie :
a) la loi de composition $\divideontimes$ est *associative* ;
b) il existe un *élément neutre* pour la loi de composition $\divideontimes$.

Exemples

1) $(\mathbb{N}, +)$ est un monoïde, car l'addition est associative et 0 est l'élément neutre pour l'addition dans $\mathbb{N}$.
2) A est un ensemble ; $\mathfrak{P}(A)$ est l'ensemble des parties de A. $(\mathfrak{P}(A), \cap)$ est un monoïde ; il en est de même de $(\mathfrak{P}(A), \cup)$.

Théorème : un monoïde possède *un seul* élément neutre ◆.

◆ Ce théorème a été démontré page 22.

Vous noterez au passage la *généralité* de ce théorème : quel que soit le monoïde, ce résultat est vrai. Ce qui importe, c'est la *structure* et non les éléments qui composent l'ensemble.
Un monoïde (E, ✲) est *commutatif*, si la loi de composition est commutative.

Exemples

1) $(\mathbb{N}, +)$ est un monoïde commutatif.
2) $(\mathfrak{P}(A), \cap)$ et $(\mathfrak{P}(A), \cup)$ sont des monoïdes commutatifs.

Théorème : (E, ✲) est un monoïde ; si l'élément a de E possède un symétrique a', alors l'élément a est *régulier* (voir page 25).

Théorème : (E, ✲) est un monoïde ; si l'élément a possède un symétrique, alors ce symétrique est *unique*.
En effet, supposons que a possède deux symétriques distincts, notés u et v.
D'après l'associativité, $u \star (a \star v) = (u \star a) \star v$
par hypothèse, $a \star v = e$ et $u \star a = e$
donc $u \star e = e \star v$
d'où $u = v$.

Exercice

Prenez $A = \{ x, y \}$ et le monoïde $(\mathfrak{P}(A), \cap)$. Existe-t-il des éléments réguliers dans ce monoïde ?
Même problème pour le monoïde $(\mathfrak{P}(A), \cup)$.

GROUPE

E est un ensemble muni d'une loi de composition notée ✲.
Le couple (E, ✲) est un *groupe* si :
a) la loi de composition ✲ est *associative ;*
b) il existe un *élément neutre* pour la loi de composition ✲ ;
c) tout élément de E possède un *symétrique* pour la loi de composition ✲.

En d'autres termes, « (E, $\divideontimes$) est un groupe » signifie :
1) (E, $\divideontimes$) est un *monoïde ;*
2) tout élément du monoïde (E, $\divideontimes$) possède un *symétrique.*
Si la loi de composition $\divideontimes$ est *commutative*, le groupe (E, $\divideontimes$) est dit *commutatif.*

Premières conséquences

Dans un groupe (E, $\divideontimes$) :
1) l'élément neutre est *unique* ;
2) tout élément est *régulier ;*
3) tout élément possède un symétrique *unique.*

Exemples et contre-exemples

1) L'ensemble des entiers $\mathbb{Z}$, muni de l'addition, est un groupe. Par contre, l'ensemble des naturels $\mathbb{N}$, muni de l'addition, n'est pas un groupe, c'est un monoïde.
2) L'ensemble $\mathbb{Z}$ muni de la multiplication n'est pas un groupe ; dans $\mathbb{Z}$, l'élément neutre pour la multiplication est 1 (voir élément neutre, ci-dessus). Le symétrique de 10 serait le nombre entier y tel que $10y = 1$. Il n'existe pas.
L'ensemble des rationnels non nuls $\mathbb{Q}^*$, muni de la multiplication, est un groupe ◆.
3) L'ensemble $\{ e, a \}$ muni de la loi de composition $\divideontimes$, définie par la table de Pythagore ci-dessous, est un groupe.

◆ Voir l'article « les Nombres ».

$\divideontimes$	e	a
e	e	a
a	a	e

Le lecteur vérifiera que la loi de composition $\divideontimes$ est associative, e est l'élément neutre ; le symétrique de e est e, le symétrique de a est a.

$\perp$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

De même, l'ensemble $\{ e, a, b \}$ muni de la loi de composition $\perp$, définie par la table de Pythagore ci-contre, est un groupe.

Définition

(E, $\divideontimes$) est un groupe ; si l'ensemble E est fini et si le cardinal de E est n, on dit que le groupe (E, $\divideontimes$) est *d'ordre n.* Si E n'est pas fini, le groupe (E, $\divideontimes$) est dit *d'ordre infini.*

Propriétés des groupes

(E, $\divideontimes$) est un groupe.

1) *x est un élément de* E, *x' son symétrique.*
Inversement, x' est le symétrique de x.
En effet, désignons par y le symétrique de x'.
Nous avons : $x' \divideontimes x = e$ et $x \divideontimes x' = e$
$y \divideontimes x' = e$ et $x' \divideontimes y = e$
Nous avons encore : $x \divideontimes x' = y \divideontimes x'$
Comme x' est régulier, il vient : $x = y$.

2) *Si u et v sont deux éléments de* E, *u' et v' leurs symétriques respectifs, le symétrique de $u \divideontimes v$ est $v' \divideontimes u'$.*
En effet, posons $a = (u \divideontimes v) \divideontimes (v' \divideontimes u')$.
D'après l'associativité, $a = u \divideontimes (v \divideontimes v') \divideontimes u'$.
Or $v \divideontimes v' = e$; donc $a = u \divideontimes e \divideontimes u'$.
Or $u \divideontimes e = u$; donc $a = u \divideontimes u'$.
Or $u \divideontimes u' = e$; d'où : $a = e$.
On démontre de même que $b = (v' \divideontimes u') \divideontimes (u \divideontimes v)$ est neutre.
Il s'ensuit que $v' \divideontimes u'$ est le symétrique de $u \divideontimes v$.

3) *Équation dans un groupe*
(E, $\divideontimes$) est un groupe ; a et b deux éléments de E.
Résoudre dans E l'équation

$$a \divideontimes x = b$$

c'est déterminer l'ensemble S des éléments de E qui, substitués à x, dans $a \divideontimes x = b$, donnent un énoncé vrai.
En désignant par a' le symétrique de a :
$a \divideontimes x = b$ est équivalent à $a' \divideontimes (a \divideontimes x) = a' \divideontimes b$
ou encore $(a' \divideontimes a) \divideontimes x = a' \divideontimes b$.
Or $a' \divideontimes a = e$ et $e \divideontimes x = x$, d'où $x = a' \divideontimes b$.
Il en résulte que :

$$a \divideontimes x = b \text{ est équivalent à } x = a' \divideontimes b$$

Ainsi, il existe un élément unique du groupe $a' \divideontimes b$ solution de l'équation.
L'ensemble S des solutions de $a \divideontimes x = b$ est le *singleton* $\{ a' \divideontimes b \}$.
On démontre de même que l'équation $x \divideontimes a = b$ admet dans E la solution unique $b \divideontimes a'$.

4) *Carré latin*. Dans la table de Pythagore d'un groupe (E, $\divideontimes$), tout élément de E figure une fois et une seule sur chaque ligne, une fois et une seule sur chaque colonne. (On dit que la table de Pythagore d'un groupe est un *carré latin*.)

Démontrons cette propriété. a est élément de E ; désignons par A l'application dans E :

$$A : x \longmapsto a \divideontimes x.$$

L'application A est une *bijection* sur E.
En effet, A est *surjective :*
quel que soit l'élément b de E, l'équation $a \divideontimes x = b$ possède une solution : $a' \divideontimes b$.
A est *injective :*
si $a \divideontimes x = a \divideontimes y$, comme a est régulier, $x = y$.
Il en résulte que, sur la ligne de l'élément a de la table de Pythagore, on trouve tous les éléments de E une fois et une seule.

En utilisant de même l'application A′ dans E :

$$A' : x \longmapsto x \divideontimes a.$$

On montre de même que, sur la colonne de l'élément a, on trouve tous les éléments de E une fois et une seule.
Cependant, inversement, si la table de Pythagore d'une loi de composition $\divideontimes$ sur un ensemble E est un carré latin, cela n'entraîne pas nécessairement que (E, $\divideontimes$) est un groupe.

Voici un exemple :

$\divideontimes$	a	b	c
a	b	a	c
b	a	c	b
c	c	b	a

Nous avons bien un carré latin, mais la loi de composition $\divideontimes$ ne confère pas à l'ensemble $\{ a, b, c \}$ la structure de groupe, car il n'existe pas d'élément neutre.

$\perp$	a	b	c
a	c	a	b
b	a	b	c
c	b	c	a

Par contre, la table de Pythagore de la loi $\perp$ définie sur le même ensemble E lui confère une structure de groupe.

b est l'élément neutre
(en effet $a \perp b = a$, etc.),
a est l'inverse de c,
c est l'inverse de a.

5) *Puissances dans un groupe*

(E, ✳) est un groupe, x un élément de E.
Posons $x^2 = x \ast x$ (carré de x)
$x^3 = x \ast x \ast x$ (cube de x)
Plus généralement, pour un naturel n

$$x^n = \underbrace{x \ast x \ast x \ast \ldots \ast x}_{n \text{ termes}} = x^{n-1} \ast x$$

est la *puissance* de x d'exposant n.

Pour $n = 1$ $\quad x^1 = x$

Pour $n = 0$ $\quad x^0 = e$

où e est l'élément neutre du groupe.

En notant alors par x^{-1} le symétrique de x, nous poserons par convention $\quad x^{-2} = x^{-1} \ast x^{-1}$

$$x^{-3} = x^{-1} \ast x^{-1} \ast x^{-1}$$

et pour un naturel n

$$x^{-n} = \underbrace{x^{-1} \ast x^{-1} \ast \ldots \ast x^{-1}}_{n \text{ termes}} = x^{-(n-1)} \ast x^{-1}$$

Il est alors facile de voir en utilisant l'*associativité* que, quels que soient les naturels m et n,

$$x^m \ast x^n = x^{m+n}$$
$$x^m \ast x^{-n} = x^{m-n}$$
$$(x^m)^n = x^{mn}$$
$$(x^{-m})^n = x^{-mn}$$

Reprenez alors le groupe $(\{e, a, b\}, \perp)$ de la page [18] ; en utilisant la table de Pythagore, il est aisé de voir que

$$a^1 = a \qquad b^1 = b$$
$$a^2 = b \qquad b^2 = a$$
$$a^3 = e \qquad b^3 = e$$

plus généralement, (E, ✳) est un groupe d'*ordre* n (E possède n éléments), x un élément de E. Considérons les puissances successives de x :

$$x^0,\ x^1,\ x^2,\ \ldots,\ x^k,\ \ldots$$

Ces puissances sont des éléments de E. Comme E est fini, il existe au moins deux naturels distincts k et h $(k > h)$ tels que $x^k = x^h$ ◆.

Or $\quad x^k \ast (x^{-h}) = x^h \ast (x^{-h})$

ou encore $\quad x^{k-h} = x^0 = e$

Il existe donc un naturel $p = k - h$, tel que $x^p = e$.

◆ Si nous reprenons l'exemple de la page 29. $E = \{e, a, b\}$ muni de la loi $\perp$, nous voyons que :
$a^1 = a$
$a^4 = (a \perp a) \perp (a \perp a) = b \perp b = a$
Nous avons aussi :
$a^8 = a^4 \perp a^4 = a \perp a = b = a^2$.

$\perp$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Le plus petit naturel non nul p, tel que $x^p = e$, s'appelle *l'ordre de l'élément* x.
Dans un groupe fini, l'ordre d'un élément est inférieur ou égal à l'ordre du groupe.
Théorème : (E, ✳) est un groupe ; x un élément de E, d'ordre p.
Les éléments

$$x^0,\ x^1,\ x^2,\ \ldots,\ x^{p-1}\ (p \text{ éléments})$$

sont tous *distincts.*
En effet, supposons qu'il existe deux naturels k et h ($k > h$, $k \leqslant p-1$ et $h < p-1$) tels que

$$x^k = x^h$$

alors

$$x^{k-h} = e$$

ou en posant $k - h = r$, $x^r = e$ avec $r < p$,
ce qui est en contradiction avec la définition de l'ordre de l'élément x.
Il est facile de voir que, si x est d'ordre p ($x^p = e$) :

$$x^{p+1} = x \quad (x^{p+1} = x^p \ast x = e \ast x)$$
$$x^{p+n} = x^n \quad (x^{p+n} = x^p \ast x^n = e \ast x^n)$$
$$x^{p-1} = x^{-1}$$

et

$$x^{-p} = e$$

Ainsi, si x est d'ordre p, son symétrique x^{-1} est lui aussi d'ordre p.
Si (E, ✳) est un groupe infini, il se peut que l'on ne puisse pas trouver un naturel non nul p tel que $x^p = e$.
On dit alors que l'ordre de x est infini.
C'est le cas pour le groupe ($\mathbb{Z}$, +) dans lequel tous les éléments (sauf 0) sont d'ordre infini.

Exemples importants de groupes

Groupes cycliques

(E, ✳) est un groupe fini d'ordre n, d'élément neutre e.
Si tous les éléments de E *sont les puissances successives d'un élément* a, (E, ✳) *est appelé groupe cyclique d'ordre* n. On le note C_n :

$$C_n = \{\ e,\ a,\ a^2,\ a^3,\ \ldots,\ a^{n-1}\ \} \text{ avec } a^n = e.$$

L'élément a est un *générateur* de ce groupe.

Exemples :

— Se reporter d'abord à la page 19 : l'exemple des rotations du disque d'angle 1, 2, 3, 4 ou 5 cinquièmes de tour nous donne un exemple de groupe cyclique d'ordre 5.
L'élément neutre est la rotation nulle.

$\ast$	e	a
e	e	a
a	a	e

Groupe cyclique d'ordre 2.

$C_2 = \{ e, a \}$

avec $a^2 = e$

$\ast$	e	a	a^2
e	e	a	a^2
a	a	a^2	e
a^2	a^2	e	a

Groupe cyclique d'ordre 3.

$C_3 = \{ e, a, a^2 \}$

avec $a^3 = e$.

On a alors $a^2 \ast a^2 = a^4 = a^3 \ast a = e \ast a = a$.

A titre d'exercice, construisez la table de Pythagore du groupe cyclique d'ordre 4 et celle du groupe cyclique d'ordre 5.
$(\mathbb{Z}, +)$ est un groupe cyclique infini.

Propriétés des groupes cycliques

1) Les groupes cycliques sont commutatifs.
Démontrons cette propriété pour C_n :
x et y sont deux éléments de C_n. Il existe deux naturels p et q tels que :

$$x = a^p \quad \text{et} \quad y = a^q$$

D'autre part,

$$x \ast y = a^p \ast a^q = a^{p+q} = a^{q+p} = a^q \ast a^p.$$

Il s'ensuit que, quels que soient les éléments x et y de C_n :

$$x \ast y = y \ast x$$

2) x est un élément de C_n ; si k est un naturel tel que $x^k = e$, alors k est un *multiple* de n.

3) a est un générateur de C_n ; l'élément a^k est un autre générateur de C_n si et seulement si k est *premier* avec n.
Exemple : dans le groupe C_3, a et a^2 sont générateurs de ce groupe.

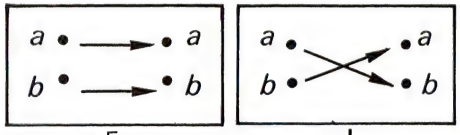
Groupes de bijections dans un ensemble

Reprenons l'exemple de la page 20 : l'ensemble $A = \{ a, b, c \}$ et l'ensemble $B = \{ E, F, G, H, I, J \}$ des bijections sur A. L'ensemble B muni de la composition des applications est un groupe.
En effet, la composition des applications est *associative*.
E est l'*élément neutre* ; tout élément de B possède un *symétrique*.

Le groupe (B, ○) est le *groupe des bijections* dans l'ensemble A. Plus généralement, A étant un ensemble à n éléments, B étant l'ensemble des bijections sur A, muni de la composition des bijections (B,○) est le groupe des bijections dans l'ensemble A, ou encore le *groupe symétrique* S_n.
L'ensemble B possède :

$$n \times (n-1) \times (n-2) \ldots \times 3 \times 2 \times 1 \text{ éléments.}$$

Ce naturel s'appelle *factorielle* n et se note n !
Pour $n = 3$, $3! = 3 \times 2 \times 1 = 6$; on est conduit au groupe S_3 déjà rencontré page [6].
Pour $n = 2$, $2! = 2$. Le groupe S_2 est d'ordre 2.



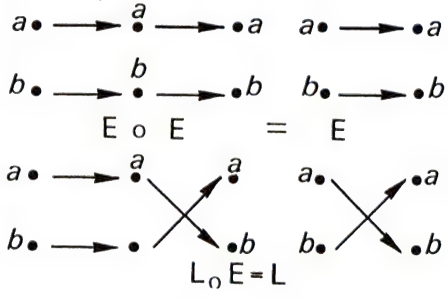
Voici les deux bijections sur $\{ a, b \}$

La table de Pythagore de ce groupe est la suivante :

○	E	L
E	E	L
L	L	E

Comparez cette table à celle du groupe cyclique d'ordre 2.

Par exemple :



Les autres cas du tableau se complètent de la même façon.

Stabilité

Revenons à la table de Pythagore du groupe symétrique S_3 (page [7]). Considérons la partie $\{ E, I, J \}$ de B muni de la loi de composition ○ que nous appelons C.

○	E	F	G	H	I	J
E	E	F	G	H	I	J
F	F	E	J	I	H	G
G	G	I	E	J	F	H
H	H	J	I	E	G	F
I	I	G	H	F	J	E
J	J	H	F	G	E	I

Table de Pythagore du groupe symétrique S_3. La partie ombrée représente les composés d'éléments de $\{ E, I, J \}$.

Nous constatons que, quel que soit le couple fait à partir de E, I ou J, le composé est E, I ou J et non pas un élément extérieur à $C = \{ E, I, J \}$ (non pas, par exemple, G ou H appartenant au groupe initial S_3).

○	E	I	J
E	E	I	J
I	I	J	E
J	J	E	I

Nous dirons que C est une *partie stable* de B pour la loi de composition ○.

Plus généralement, (A, $\divideontimes$) étant un groupe, A′ une partie de A telle que, quel que soit le couple (x, y) élément de A′ $\times$ A′, le composé $x \divideontimes y$ est élément de A′. A′ *est une partie stable de* A *pour la loi de composition* $\divideontimes$.

En général, une partie d'un groupe n'est pas stable pour la loi du groupe. Si nous reprenons le tableau précédent, nous pouvons extraire des parties non stables pour la loi ○.

Dans le cas où A′ est une partie stable de A, nous pouvons définir dans A′ la loi de composition notée $\perp$, telle que, pour tout couple (x, y) élément de A′ $\times$ A′, on ait

$$x \perp y = x \divideontimes y$$

Cette loi de composition $\perp$ est la *loi de composition induite* dans A′ par la loi de composition $\divideontimes$ dans A.

En principe, il ne faudrait pas confondre les deux symboles $\perp$ et $\divideontimes$, car $\perp$ est une loi de composition dans A′ et $\divideontimes$ une loi de composition dans A. Cependant, dans la pratique, on utilisera le même symbole pour désigner les deux lois de composition. En utilisant la table de Pythagore du groupe symétrique S_3, il est facile de vérifier

que la partie $M = \{ E, F, I \}$ de B *n'est pas stable* pour la loi de composition $\circ$:

$\circ$	E	F	I
E	E	F	I
F	F	E	H
I	I	G	J

En effet, on a, par exemple : $F \circ I = H$, et H n'est pas élément de M.

Sous-groupe

(E, $\ast$) est un groupe et H une partie de E ; si H munie de la loi induite $\ast$ a une structure de groupe (associativité, élément neutre, symétrique), nous dirons que H est un sous-groupe de E.
L'ensemble des entiers $\mathbb{Z}$ a une structure de groupe pour l'addition. Considérons le sous-ensemble $\mathbb{P}$ des entiers pairs : il a une structure de groupe pour l'addition (la loi + est associative, l'élément neutre 0 appartient à $\mathbb{P}$ et chaque élément de $\mathbb{P}$ admet un symétrique pour l'addition dans $\mathbb{P}$: $4 + (-4) = 0$). $\mathbb{P}$ est un *sous-groupe* de $\mathbb{Z}$. Mais l'ensemble $\mathbb{N}$ des naturels n'a pas une structure de groupe pour l'addition (la loi + est associative, l'élément neutre appartient à l'ensemble $\mathbb{N}$, mais les éléments de $\mathbb{N}$ n'ont pas de symétrique). C'est un monoïde et non un sous-groupe de $\mathbb{Z}$.

Exemples

1) Le groupe (E, $\ast$) possède au moins deux sous-groupes (E, $\times$) et ($\{ e \}$, $\ast$).
2) Considérons le groupe cyclique $C_4 = \{ e, a, a^2, a^3 \}$ avec $a^4 = e$. Voici la table de Pythagore de ce groupe :

$\ast$	e	a	a^2	a^3
e	e	a	a^2	a^3
a	a	a^2	a^3	e
a^2	a^2	a^3	e	a
a^3	a^3	e	a	a^2

Prenons la partie $H = \{ e, a^2 \}$ de C_4. Cette partie est stable :

$\ast$	e	a^2
e	e	a^2
a^2	a^2	e

En outre, il est facile de vérifier que (H, $\ast$) est un groupe. (H, $\ast$) est donc un sous-groupe de C_4.
Il est important de noter que si H est une partie stable d'un groupe (E, $\ast$), alors (H, $\ast$) n'est pas nécessairement un sous-groupe de (E, $\ast$). Ainsi, par exemple, $\mathbb{N}$ est une partie stable du groupe

$(\mathbb{Z}, +)$, cependant $(\mathbb{N}, +)$ n'est pas un sous-groupe de $(\mathbb{Z}, +)$, car $(\mathbb{N}, +)$ n'est pas un groupe ; c'est un monoïde.

Théorème : (E, ✳) est un groupe ; e désigne son élément neutre.

Si (H, ✳) est un sous-groupe de (E, ✳), alors :

a) e est l'élément neutre de (H, ✳) ;

b) u est un élément de H, u' est le symétrique de u dans (E, ✳), u' est aussi le symétrique de u dans (H, ✳).

En effet, quel que soit l'élément x de E,

$$x \ast e = x \qquad \text{et} \qquad e \ast x = x$$

En particulier, ces deux égalités sont vraies quel que soit l'élément x de H ; donc e est l'élément neutre du groupe (H, ✳). D'autre part, $u \ast u' = e$ et $u' \ast u = e$ car u et u' sont symétriques dans (E, ✳). e étant aussi l'élément neutre de (H, ✳), on en déduit que u et u' sont aussi symétriques dans (H, ✳).

Caractérisation d'un sous-groupe

(E, ✳) est un groupe ; e désigne son élément neutre.

H est une *partie* de E.

(H, ✳) est un *sous-groupe* de (G, ✳), *si et seulement si :*

1) H est *une partie stable* de E pour la loi de composition ✳ ;

2) l'*élément neutre* e de (E, ✳) est *élément* de H ;

3) quel que soit l'élément x de H, son *symétrique* x' dans (E, ✳) est aussi élément de H.

En effet, si (H, ✳) est un sous-groupe de (E, ✳) :

1) H est une partie stable de (E, ✳) ;

2) l'élément neutre e de (E, ✳) est aussi élément neutre de (H, ✳), donc $e \in H$;

3) x' est aussi le symétrique de x dans (H, ✳), donc $x' \in H$.

Réciproquement, si H est une partie de E, vérifiant les conditions 1, 2 et 3, alors :

1) e est élément neutre de (H, ✳) ;

2) x' est le symétrique de x dans (H, ✳) ;

3) l'associativité est une conséquence de la stabilité ; en effet, quels que soient les éléments x, y, z de H, nous avons

$$(x \ast y) \ast z = x \ast (y \ast z)$$

car cela est vrai dans (E, ✳).
Il s'ensuit que (H, ✳) est un groupe, donc un sous-groupe de (E, ✳).

Application : Trouver les sous-groupes du groupe cyclique C_4 qui contiennent a^3.

D'après l'étude précédente, un tel sous-groupe doit contenir e et le symétrique de a^3, soit a. Comme un sous-groupe est une partie stable de C_4, il doit contenir $a \divideontimes a$, c'est-à-dire a^2. Le seul sous-groupe de C_4 qui contienne a^3 est C_4 lui-même.

Théorème : (E, $\divideontimes$) est un groupe fini d'ordre n ; e désigne son élément neutre. a est un élément de E, d'ordre p : $a^p = e$.
Si H désigne l'ensemble des puissances de a :

$$H = \{ a, a^2, a^3, ..., a^p \}$$

alors (H, $\divideontimes$) est un sous-groupe de (E, $\divideontimes$).
(H, $\divideontimes$) est le *sous-groupe cyclique* de (E, $\divideontimes$) *engendré* par l'élément a, on le note $< a >$.
En effet, H est stable pour la loi de composition $\divideontimes$.
Comme $a^p = e$, l'élément neutre de E est élément de H. Nous savons d'autre part que H est le groupe cyclique d'ordre p : chaque élément de H admet donc un symétrique dans H.
(H, $\divideontimes$) est bien un sous-groupe de (E, $\divideontimes$).

Exercice : Trouver le sous-groupe cyclique $< a^2 >$ de C_6.

Autre caractérisation d'un sous-groupe :
(E, $\divideontimes$) est un groupe ; e désigne son élément neutre.
H étant *une partie non vide* de E, (H, $\divideontimes$) est un sous-groupe de (E, $\divideontimes$) *si et seulement si*, pour tout couple (x, y) de $H \times H$, $x \divideontimes y'$ est élément de H, où y' désigne le symétrique de y dans (E, $\divideontimes$).
En effet, si (H, $\divideontimes$) est un sous-groupe de (E, $\divideontimes$), alors quels que soient x et y éléments de H, y' est aussi élément de H et, compte tenu de la stabilité, $x \divideontimes y'$ est aussi élément de H.
Réciproquement, soit x un élément de H ; appliquons la condition au couple (x, x) de $H \times H$:

$$x \divideontimes x' \in H$$

or $\quad x \divideontimes x' = e, \quad$ donc $\quad e \in H.$

Appliquons la condition au couple (e, x) de $H \times H$:

$$e \divideontimes x' \in H$$

or $\quad e \divideontimes x' = x', \quad$ donc $\quad x' \in H.$

Ainsi, quel que soit x, élément de H, son symétrique x' est aussi élément de H.
x et y sont deux éléments de H. Nous venons de voir que y' est aussi élément de H. Appliquons la conditions au couple (x, y') de $H \times H$:

$$x \divideontimes (y')' \in H$$

or $\quad (y')' = y, \quad$ donc $\quad x \divideontimes y \in H.$

Il s'ensuit que H est une partie stable de (E, $\divideontimes$).
(H, $\divideontimes$) est un groupe, donc un sous-groupe de (E, $\divideontimes$).

Morphismes de groupes

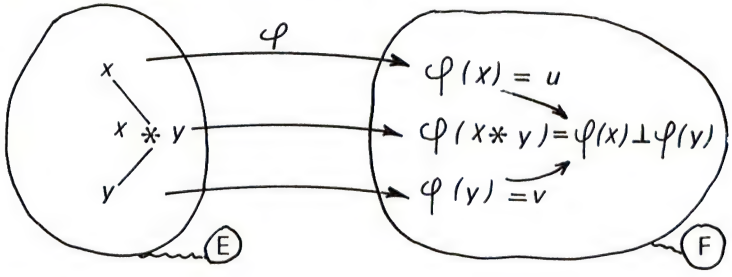
Isomorphisme

En algèbre, nous nous intéressons beaucoup plus aux lois de composition qui confèrent à un ensemble une structure et à la structure elle-même qu'aux éléments de cet ensemble.
Et, si nous étudions une application d'un groupe dans un autre, il est intéressant de savoir si cette application respecte les lois de composition des deux groupes.
Si (E, $\divideontimes$) et (F, $\perp$) sont deux groupes, si φ est une *application* de *source* E et de *but* F, « l'application φ est un *isomorphisme* de (E, $\divideontimes$) sur (F, $\perp$) » signifie :

> 1) φ est une bijection de E sur F.
> 2) Quel que soit (x, y) élément de $E \times E$:
> $$\varphi(x \divideontimes y) = \varphi(x) \perp \varphi(y)$$

Autrement dit, la bijection φ « respecte » les lois de composition.
(E, $\divideontimes$) est *isomorphe* à (F, $\perp$).
Nous noterons dans ce cas (E, $\divideontimes$) $\sim$ (F, $\perp$) ◆.

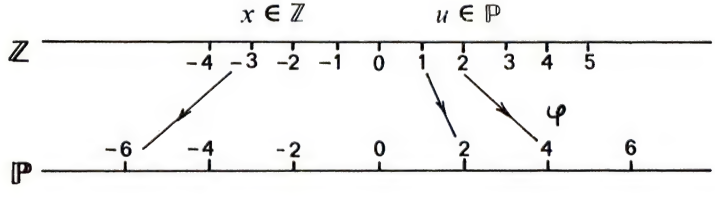
◆ $\sim$ = isomorphe.



Exemple

Nous avons vu page 37 que l'ensemble des entiers $\mathbb{Z}$ et l'ensemble $\mathbb{P}$ des entiers pairs ont une structure de groupe pour l'addition. Nous pouvons définir un *isomorphisme* entre ces deux groupes. A tout élément x de $\mathbb{Z}$ nous associons par φ son double $u = 2x$ élément de $\mathbb{P}$:

$$x \longmapsto \varphi(x) = u = 2x.$$



1) Le lecteur pourra démontrer que φ est une bijection de $\mathbb{Z}$ dans $\mathbb{P}$ (on peut se reporter à un exemple analogue traité dans l'article « les Ensembles », page 268).

2) Pour tout x et y de $\mathbb{Z}$, nous avons la propriété :

$$\varphi(x + y) = \varphi(x) + \varphi(y).$$

Ainsi, par exemple, prenons $x = 3$ et $y = 4$.

$$x + y = 7 \quad \text{et} \quad \varphi(x + y) = \varphi(7) = 14.$$

D'autre part, $\varphi(x) = \varphi(3) = 6 \quad \varphi(y) = \varphi(4) = 8$.

Or $14 = 6 + 8$, donc $\varphi(x + y) = \varphi(x) + \varphi(y)$.

L'isomorphisme est une relation entre groupes. Cette relation est symétrique ; en effet, si φ est un isomorphisme de $(E, \divideontimes)$ sur $(F, \perp)$, en posant

$$\varphi(x) = u \quad \text{avec} \quad u = \varphi^{-1}(x) \blacklozenge.$$

$$\varphi(y) = v \quad \text{avec} \quad v = \varphi^{-1}(y)$$

◆ Nous rappelons qu'une bijection d'un ensemble E vers un ensemble F, notée f, associe à tout x de E son image u dans F ($u = f(x)$) et *réciproquement* permet d'associer à u dans F son image x dans E par la bijection réciproque notée f^{-1} ($x = f^{-1}(u)$).

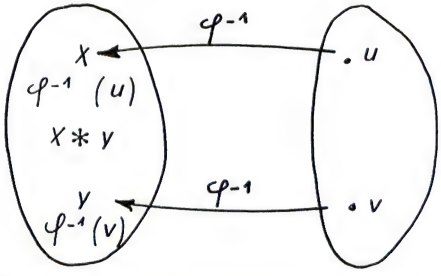
où φ^{-1} est la réciproque de φ,

nous avons

$$\varphi^{-1}(u \perp v) = \varphi^{-1}[\varphi(x) \perp \varphi(y)] = \varphi^{-1}[\varphi(x \divideontimes y)] = x \divideontimes y$$

d'où

$$\varphi^{-1}(u \perp v) = \varphi(u) \divideontimes \varphi(v).$$



Il en résulte que φ^{-1} est un isomorphisme de $(F, \perp)$ sur $(E, \divideontimes)$. Il est facile de montrer que l'isomorphisme est une relation d'équivalence sur les groupes. Une classe d'équivalence s'appelle parfois un « groupe abstrait » et un élément d'une classe est un « modèle » de ce groupe abstrait.

Théorème de Cayley

Tout groupe fini $(G, \divideontimes)$ d'ordre n est isomorphe à un sous-groupe du groupe symétrique $\mathcal{S}_n$. Nous allons démontrer qu'il existe un sous-groupe $(H, \bullet)$ de $\mathcal{S}_n$ isomorphe à $(G, \divideontimes)$.

a désigne un élément de G.
h_a est l'application de G vers G définie par

$$h_a : x \longmapsto x \divideontimes a$$

Cette application est une bijection de G sur G, h_a est donc un élément de $\mathcal{S}_n$.
A chaque élément de G correspond ainsi une bijection de G sur G.
Désignons par H l'ensemble de ces bijections :

$$H = \{ h_a,\ a \in G \}$$

(H, $\bullet$) est un sous-groupe de $\mathcal{S}_n$:

a) H est une partie stable de $\mathcal{S}_n$.
Soit h_a et h_b deux éléments de H ; la bijection composée $h_a \bullet h_b$ est

$$h_a \bullet h_b : x \longmapsto (x \divideontimes a) \divideontimes b$$

or $$(x \divideontimes a) \divideontimes b = x \divideontimes (a \divideontimes b)$$

Il en résulte que $$h_a \bullet h_b = h_{a \divideontimes b}$$

Comme $a \divideontimes b$ est élément de G, $h_{a \divideontimes b}$ est élément de H.

b) e désigne l'élément neutre de (G, $\divideontimes$).
La bijection $$h_e : x \longmapsto x \divideontimes e$$

ou encore $$h_e : x \longmapsto x$$

est la bijection identique sur G.
C'est l'élément neutre de (H, $\bullet$). En effet, quel que soit h_a élément de H

$$h_a \bullet h_e = h_{a \divideontimes e} = h_a$$

et $$h_e \bullet h_a = h_{e \divideontimes a} = h_a$$

c) Quels que soient h_a et h_b éléments de H, $h_a \bullet h_b = h_{a \divideontimes b}$; en particulier, pour $b = a^{-1}$ (a^{-1} est le symétrique de a dans G), nous avons : $$h_a \bullet h_{a^{-1}} = h_{a \divideontimes a^{-1}} = h_e$$

Tout élément h_a de H admet un symétrique $h_{a^{-1}}$.
(H, $\bullet$) est donc un sous-groupe d'ordre n de $\mathcal{S}_n$.
Il reste à établir que (G, $\divideontimes$) est isomorphe à (H, $\bullet$).
Désignons par φ la bijection de G vers H telle que

$$\varphi : a \longmapsto h_a$$

Nous avons, quels que soient les éléments a et b de G :

$$\varphi(a \divideontimes b) = h_{a \divideontimes b}$$

$$\varphi(a) \bullet \varphi(b) = h_a \bullet h_b = h_{a \divideontimes b}$$

donc $$\varphi(a \divideontimes b) = \varphi(a) \bullet \varphi(b)$$

φ est donc un isomorphisme de (G, $\divideontimes$) vers (H, $\bullet$).

Homomorphisme

Nous venons de voir que l'application φ est un isomorphisme de $(E, \divideontimes)$ sur $(F, \perp)$ si φ est une bijection de E sur F et si φ respecte les lois de composition.

Si nous abandonnons la première condition (φ est une bijection), nous dirons que φ est un *homomorphisme* du groupe $(E, \divideontimes)$ vers $(F, \perp)$.

En d'autres termes :

L'*application* $\varphi : E \longrightarrow F$ est un *homomorphisme* si quel que soit l'élément (x, y) de $E \times E$

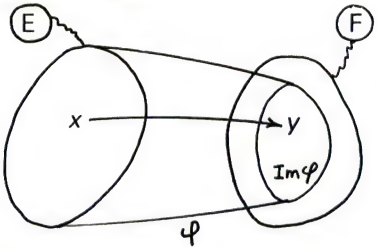
$$\varphi(x \divideontimes y) = \varphi(x) \perp \varphi(y) \blacklozenge.$$

◆ Propriété 2 de l'isomorphisme (voir page 40).

Un homomorphisme *surjectif* est un *épimorphisme.*

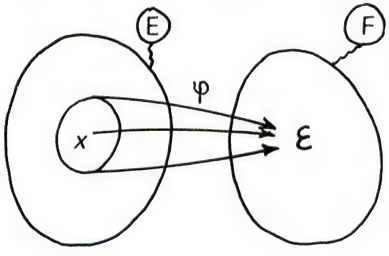
Un homomorphisme *injectif* est un *monomorphisme.*

Bien entendu, un homomorphisme *bijectif* est un *isomorphisme.* $(E, \divideontimes)$ et $(F, \perp)$ sont deux groupes et $\varphi : E \longrightarrow F$ est un homomorphisme.



On appelle *image* par φ le sous-ensemble de F dont les éléments sont de la forme $\varphi(x)$, où x décrit E.
Cette partie de F se note Imφ ou φ(E).

Désignons par ε l'élément neutre de $(F, \perp)$.



On appelle *noyau* de l'homomorphisme φ le sous-ensemble de E dont les éléments ont pour image ε :

$$f(x) = \varepsilon$$

Cette partie de E se note Ker φ ◆.

◆ « Noyau » se dit « kernel » en anglais.

Théorème

L'élément neutre e de $(E, \divideontimes)$ est élément du noyau. En effet, quel que soit l'élément x de E :

$$\varphi(x) = \varphi(e \divideontimes x) = \varphi(e) \perp \varphi(x)$$

Comme $\varphi(x)$ est régulier, il vient $\varphi(e) = \varepsilon$
donc $e \in$ Ker φ.

Théorème

(Imφ, $\perp$) est un sous-groupe de (F, $\perp$) . En effet, soit $\varphi(x)$ et $\varphi(y)$ deux éléments de Imφ :

$$\varphi(x) \perp \varphi(y) = \varphi(x \divideontimes y)$$

Comme $\varphi(x \divideontimes y)$ est élément de Imφ, il en est de même de $\varphi(x) \perp \varphi(y)$. Im$\varphi$ est stable.

La loi de composition $\perp$ est associative sur Imφ.
ε est élément neutre de (Imφ, $\perp$).

Quel que soit $\varphi(x)$ élément de Imφ :

$$\varepsilon = \varphi(e) = \varphi(x \divideontimes x^{-1}) = \varphi(x) \perp \varphi(x^{-1})$$

et
$$\varepsilon = \varphi(e) = \varphi(x^{-1} \divideontimes x) = \varphi(x^{-1}) \perp \varphi(x).$$

Il en résulte que $\varphi(x)$ et $\varphi(x^{-1})$ sont symétriques dans Imφ.

D'où le théorème.

Théorème

(Ker φ, $\divideontimes$) est un sous-groupe de (E, $\divideontimes$).

En effet, soit x et y deux éléments de Ker φ.

$$\varphi(x) = \varepsilon \quad \text{et} \quad \varphi(y) = \varepsilon$$
$$\varphi(x) \perp \varphi(y) = \varphi(x \divideontimes y) = \varepsilon \perp \varepsilon = \varepsilon$$

donc
$$x \divideontimes y \in \text{Ker } \varphi$$

Ker φ est stable.

La loi de composition $\divideontimes$ est associative sur Ker φ.
e est l'élément neutre de (Ker φ, $\divideontimes$), car $\varphi(e) = \varepsilon$.
En outre, si $\varphi(x) = \varepsilon$, nous avons aussi $\varphi(x^{-1}) = \varepsilon$.

D'où le théorème.

Groupe opérant sur un ensemble

1) *Définition*

(G, $\divideontimes$) est un groupe, d'élément neutre e. E est un ensemble. On dit que le groupe G opère sur E, s'il existe une application

$$\varphi : G \times E \longrightarrow E$$
$$(\alpha, x) \longmapsto \alpha \,.\, x$$

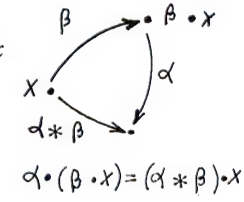
telle que

(1) Quel que soit le couple (α, β) de $G \times G$, et quel que soit x élément de E,

$$\alpha \,.\, (\beta \,.\, x) = (\alpha \divideontimes \beta) \,.\, x$$

(2) Quel que soit x élément de E

$$e \,.\, x = x$$



2) *Equivalence de transitivité*

(G, $\divideontimes$) est un groupe qui opère sur l'ensemble E.
Désignons par $\mathcal{C}$ la relation sur E, telle que

$x \mathcal{C} y$, si et seulement s'il existe un élément α de G tel que

$$x = \alpha \,.\, y.$$

Démontrons que $\mathcal{C}$ est une relation d'équivalence sur E :

a) e étant l'élément neutre de G

quel que soit x, élément de E

$$x = e \,.\, x$$

donc $$x \mathcal{C} x$$

La relation $\mathcal{C}$ est réflexive.

b) Si $x \mathcal{C} y$, il existe α élément de G tel que $x = \alpha \,.\, y$.

En désignant par α^{-1} le symétrique de α

$$\alpha^{-1} \,.\, x = \alpha^{-1} \,.\, (\alpha \,.\, y) = (\alpha^{-1} \divideontimes \alpha) \,.\, y = e \,.\, y = y$$

donc $y \mathcal{C} x$.

La relation $\mathcal{C}$ est symétrique.

c) Si $x \mathcal{C} y$ et $y \mathcal{C} z$, il existe α et β élément de G tel que

$$x = \alpha \,.\, y \qquad \text{et} \qquad y = \beta \,.\, z$$

donc $$x = \alpha \,.\, (\beta \,.\, z) = (\alpha \divideontimes \beta) \,.\, z$$

et $$x \mathcal{C} z.$$

La relation $\mathcal{C}$ est transitive.

La relation $\mathcal{C}$ s'appelle *équivalence de transitivité.*

Pour un élément x de E, désignons par $\mathcal{O}_x$ la classe de x modulo $\mathcal{C}$.
$\mathcal{O}_x$ s'appelle *l'orbite* de x.

Par définition

$$y \in \mathcal{O}_x \text{ si et seulement si } y \mathcal{C} x$$

ou encore

il existe α élément de G, tel que $y = \alpha \,.\, x$.

3) *Définitions :*

(G, $\divideontimes$) est un groupe qui opère sur l'ensemble E.

a) Si E est *la seule* orbite, on dit que le groupe (G, $\divideontimes$) opère *transitivement* sur E. Le couple (E, G) est un *espace homogène.*

b) Si (E, G) est un espace homogène, on dit que le groupe (G, $\divideontimes$) opère *fidèlement* sur E si

quels que soient x et y éléments de E, il existe α élément de G, tel que

$$x = \alpha \,.\, y.$$

ANNEAU

Nous allons maintenant envisager un ensemble E non pas muni d'une loi de composition comme le groupe mais de deux lois de composition, notées respectivement $\divideontimes$ et $\perp$.

Définition

Le triplet $(E, \divideontimes, \perp)$ est un anneau si :

a) le couple $(E, \divideontimes)$ est un groupe commutatif ;

b) la loi de composition $\perp$ est *associative* ;

c) la loi de composition $\perp$ est *distributive* sur la loi de composition $\divideontimes$.

Exemples

1) L'ensemble des entiers $\mathbb{Z}$ muni de l'addition et de la multiplication est un anneau. $\mathbb{Z}$ est un groupe commutatif pour l'addition ; la multiplication est associative et distributive par rapport à l'addition ◆.

◆ Voir le début de l'article pour les définitions de : associativité, commutativité, distributivité.

2) L'ensemble $E = \{ a \}$ muni des lois de composition $\divideontimes$ et $\perp$ telles que

$$a \divideontimes a = a \quad \text{et} \quad a \perp a = a$$

est un anneau.

3) $a \,.\, \mathbb{Z}$ désigne l'ensemble des multiples de a.

$(a \,.\, \mathbb{Z}, +, \times)$ est un anneau.

Propriétés élémentaires des anneaux

Théorème

> $(E, \divideontimes, \perp)$ est un anneau ; e est l'élément neutre de $(E, \divideontimes)$.
> On a, quel que soit x élément de E :
>
> $$x \perp e = e \perp x = e$$

On dit que e est absorbant pour la loi de composition $\perp$. En effet, quel que soit x

$$\begin{aligned} x \perp x &= x \perp (x \divideontimes e) \\ &= (x \perp x) \divideontimes (x \perp e) \end{aligned}$$

Or $x \perp x$ est régulier ; il en résulte que :

$$e = x \perp e$$

On démontre de même que $e \perp x = e$.

Dans l'anneau $(\mathbb{Z}, +, \times)$, ce résultat conduit à :

quel que soit x, $x \times 0 = 0 \times x = 0$.

Désignons par sym x le symétrique de x pour la loi de composition $\divideontimes$.

Théorème

> Quels que soient les éléments x et y de E :
> $$x \perp \text{sym } y = \text{sym } (x \perp y)$$

En effet,

$$(x \perp y) \divideontimes (x \perp \text{sym } y) = x \perp (y \divideontimes \text{sym } y)$$
$$= x \perp e = e$$

Il en résulte que

$$\text{sym } (x \perp y) = x \perp \text{sym } y$$

On démontre de même que

$$\text{sym } (x \perp y) = \text{sym } x \perp y$$

En particulier,

$$\text{sym } x \perp \text{sym } y = x \perp y$$

En effet,

$$\text{sym } x \perp \text{sym } y = \text{sym } (x \perp \text{sym } y)$$
$$= \text{sym } [\text{sym } (x \perp y)]$$

Or $$\text{sym } [\text{sym } (x \perp y)] = x \perp y$$

Dans l'anneau $(\mathbb{Z}, +, \times)$, ces résultats donnent, quels que soient les entiers x et y,

$$x \times (-y) = -(x \times y)$$

et $$(-x) \times (-y) = x \times y$$

L'anneau (E, $\divideontimes$, $\perp$) est unitaire, s'il existe un élément neutre ε pour la deuxième loi de composition $\perp$. ε est appelé élément *unité*.

Ainsi, l'anneau $(\mathbb{Z}, +, \times)$ est unitaire, car 1 est neutre pour la multiplication.

Dans un anneau unitaire (E, $\divideontimes$, $\perp$) d'élément unité ε, un élément a de E est dit *inversible* s'il existe un élément a' de E, tel que

$$a \perp a' = \varepsilon \quad \text{et} \quad a' \perp a = \varepsilon$$

L'anneau des entiers modulo p

Il est facile de voir que cette relation $\mathcal{R}$ est aussi *stable* pour la multiplication dans $\mathbb{Z}$ et il est intéressant d'étudier la loi de composition induite $\otimes$ sur C_p.
A titre d'exemple, voir ci-contre la table de multiplication dans C_6. La multiplication $\otimes$ est *associative*, *commutative*. [1] est élément unité, mais les éléments différents de [0] n'ont pas nécessairement de symétrique dans C_p pour la multiplication. Par exemple, dans C_6, [2], [3] et [4] n'ont pas de symétrique. En outre, la multiplication $\otimes$ est *distributive* sur l'addition $\oplus$. Ainsi, $(C_p, \oplus, \otimes)$ est un *anneau commutatif unitaire*.

$\otimes$	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

Anneau intègre

Dans l'anneau des entiers $(\mathbb{Z}, +, \times)$, le produit de deux entiers non nuls est un entier non nul :

$$\text{si } m \neq 0 \text{ et } n \neq 0, \text{ alors } m \times n \neq 0.$$

Plus généralement, l'anneau $(A, \divideontimes, \perp)$ est dit *intègre*, si quels que soient les éléments a et b, différents de l'élément neutre e du groupe $(A, \divideontimes)$, alors $a \perp b$ est différent de e.
Inversement, $(A, \divideontimes, \perp)$ étant un anneau, si $a \neq e$ et $a \perp b = e$ pour un élément b différent de e, alors a est appelé un *diviseur de zéro*.

Exemple

Dans l'anneau $(C_6, \oplus, \otimes)$ nous avons

$$[3] \otimes [2] = [0] \quad [3] \otimes [4] = [0]$$

$\mathbb{Z}$ désigne l'ensemble des *entiers*. $\mathcal{R}$ est la relation d'équivalence ◆ sur $\mathbb{Z}$ définie par : $x_1 \mathcal{R} x_2$ si et seulement si $x_1 - x_2$ est un *multiple* de 6. $\mathbb{Z}/\mathcal{R}$ possède six classes notées [0], [1], [2], [3], [4], et [5] ◆.

◆ Voir l'article « les Ensembles », page 281.

◆ [0] représente la classe contenant les éléments : 0, 6, 12, 18, etc.
[1] les éléments 1, 7, 13, etc.
[4] les éléments 4, 10, 16, etc.

La relation $\mathcal{R}$ est stable pour l'addition dans $\mathbb{Z}$.
Voici la table de Pythagore de la loi de composition induite $\oplus$:

$\oplus$	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[1]	[2]	[3]	[4]	[5]
[1]	[1]	[2]	[3]	[4]	[5]	[0]
[2]	[2]	[3]	[4]	[5]	[0]	[1]
[3]	[3]	[4]	[5]	[0]	[1]	[2]
[4]	[4]	[5]	[0]	[1]	[2]	[3]
[5]	[5]	[0]	[1]	[2]	[3]	[4]

Plus généralement, si $\mathcal{R}$ désigne la relation d'équivalence sur $\mathbb{Z}$ définie par : $x_1 \mathcal{R} x_2$ si et seulement si $x_1 - x_2$ est un *multiple* de p, alors $\mathbb{Z}/\mathcal{R}$ possède p classes ; $\mathcal{R}$ est stable pour l'addition dans $\mathbb{Z}$ et la loi de composition induite $\oplus$ confère à $\mathbb{Z}/\mathcal{R}$ la structure de groupe. $(\mathbb{Z}/\mathcal{R}, \oplus)$ est un groupe cyclique d'ordre p. Ce groupe est souvent noté C_p.

Il est facile de voir que $(\mathbb{Z}/\mathcal{R}, \oplus)$ est un groupe cyclique d'ordre 6.

Les éléments [2], [3] et [4] sont des diviseurs de zéro.
A titre d'exercice, le lecteur démontrera que dans un anneau intègre (A, $\divideontimes$, $\perp$)

$$\text{si} \quad a \perp b = a \perp c \quad \text{et} \quad a \neq c, \quad \text{alors} \quad b = c.$$

Sous-anneau

(A, $\divideontimes$, $\perp$) est un anneau, B une partie *stable* pour chacune des lois de composition $\divideontimes$ et $\perp$.
Si (B, $\divideontimes$, $\perp$) est un anneau, alors (B, $\divideontimes$, $\perp$) est un *sous-anneau* de (A, $\divideontimes$, $\perp$).

Exemples

1) $(a\mathbb{Z}, +, \times)$ est un sous-anneau de l'anneau des entiers $(\mathbb{Z}, +, \times)$.

2) (A, $\divideontimes$, $\perp$) est un anneau, e l'élément neutre de (A, $\divideontimes$).

$(\{ e \}, \divideontimes, \perp)$ et (A, $\divideontimes$, $\perp$) sont des sous-anneaux de (A, $\divideontimes$, $\perp$).

Morphisme d'anneaux

Isomorphisme

(A, $\bullet$, T) et (B, $\divideontimes$, $\perp$) sont deux anneaux.
φ est une *application* de *source* A et de *but* B.
« L'application φ est un *isomorphisme* de (A, $\bullet$, T) sur (B, $\divideontimes$, $\perp$) » signifie :

1) φ est une *bijection* de A sur B.
2) Quel que soit (x, y) élément de $A \times A$:

$$\varphi (x \bullet y) = \varphi (x) \divideontimes \varphi (y)$$
$$\varphi (x \ \mathrm{T} \ y) = \varphi (x) \perp \varphi (y)$$

L'anneau (A, $\bullet$, T) est *isomorphe* à l'anneau (B, $\divideontimes$, $\perp$).
L'isomorphisme est une relation entre anneaux. Cette relation est une équivalence.

Exemple

L'isomorphisme de groupe φ que nous avons défini page [46] n'est pas un isomorphisme d'anneau. En effet, nous avons bien :

1) $\mathbb{Z}$ et $\mathbb{P}$ sont des anneaux pour les lois $+$ et $\times$.

2) φ est une bijection de $\mathbb{Z}$ dans $\mathbb{P}$.

3) $\varphi (x + y) = \varphi (x) + \varphi (y)$.

Mais nous n'avons pas $\varphi (x \times y) = \varphi (x) \times \varphi (y)$.

Reprenons $x = 3$ et $y = 4$.

$\varphi(x) = 6$, $\varphi(y) = 8$ et $\varphi(x) \times \varphi(y) = 6 \times 8 = 48$, alors que $x \times y = 3 \times 4 = 12$, d'où $\varphi(x \times y) = \varphi(12) = 24$.

Or $48 \neq 24$.

Homomorphisme

« L'application φ de source A et de but B est un homomorphisme de l'anneau (A, $\bullet$, T) vers l'anneau (B, $\divideontimes$, $\perp$) » signifie quel que soit (x, y) élément de $A \times A$:

$$\varphi(x \bullet y) = \varphi(x) \divideontimes \varphi(y)$$
$$\varphi(x \mathrm{T} y) = \varphi(x) \perp \varphi(y)$$

Comme dans l'homomorphisme de groupes, nous n'avons plus la propriété de bijection.

Exemple

(A, $\divideontimes$, $\perp$) est un anneau unitaire, ε désigne l'unité de cet anneau et sym ε le symétrique de ε pour la loi de composition $\divideontimes$.

Posons :

$0\varepsilon = e$, où e est l'élément neutre du groupe (A, $\divideontimes$)

$1\varepsilon = \varepsilon$

$2\varepsilon = \varepsilon \divideontimes \varepsilon$

$3\varepsilon = 2\varepsilon \divideontimes \varepsilon$

.....

$n\varepsilon = (n - 1)\varepsilon \divideontimes \varepsilon$, ou n est un naturel.

De même :

$(-1)\varepsilon = \text{sym } \varepsilon$

$(-2)\varepsilon = \text{sym } \varepsilon \divideontimes \text{sym } \varepsilon$

$(-3)\varepsilon = (-2)\varepsilon \divideontimes \text{sym } \varepsilon$

.....

$(-n)\varepsilon = (-n + 1) \divideontimes \text{sym } \varepsilon$

Nous donnons ainsi un sens à l'écriture $z\varepsilon$ où z est un entier. Désignons par B l'ensemble des éléments de A de la forme $z\varepsilon$ où $z \in \mathbb{Z}$.

Le lecteur vérifiera à titre d'exercice que (B, $\divideontimes$, $\perp$) est un sous-anneau de (A, $\divideontimes$, $\perp$) et démontrera que dans le cas de l'anneau des entiers modulo p (C_p, $\oplus$, $\otimes$) ce sous-anneau est égal à (C_p, $\oplus$, $\otimes$).

Dans ce dernier exemple

$$p[1] = [0]$$

$$(p + 1)\ [1] = [1] \text{ etc.}$$

Notez que si (A, $\divideontimes$, $\perp$) est un anneau unitaire tel que quel que soit l'entier z non nul, $z\,\varepsilon \neq e$, où e est l'élément neutre du groupe (A, $\divideontimes$), alors le sous-anneau (B, $\divideontimes$, $\perp$) est *infini*. En effet, si pour deux éléments z_1 et z_2 de $\mathbb{Z}$

$$z_1\,\varepsilon = z_2\,\varepsilon \qquad \text{avec } z_1 \neq z_2$$

alors $\quad (z_1 \times \text{sym } z_2)\,\varepsilon = e$

ce qui contredit l'hypothèse, quel que soit z élément de $\mathbb{Z}^*$, $z\,\varepsilon \neq e$. Pour tout anneau (A, $\divideontimes$, $\perp$), l'application φ de source $\mathbb{Z}$ et de but A, telle que

$$\varphi : z \longmapsto z\,\varepsilon$$

est un *homomorphisme*.

Théorème

(A, $\divideontimes$, $\perp$) est un anneau unitaire.
Il existe un sous-anneau (B, $\divideontimes$, $\perp$) de cet anneau, isomorphe soit à l'anneau des entiers ($\mathbb{Z}$, $+$, $\times$), soit à un anneau des entiers (modulo p), ($p \geqslant 2$).

Preuve

(B, $\divideontimes$, $\perp$) désigne le sous-anneau de (A, $\divideontimes$, $\perp$) dont les éléments sont de la forme $z\,\varepsilon$ où $z \in \mathbb{Z}$.

a) Supposons que quel que soit z non nul, $z\,\varepsilon \neq e$.

L'application $\varphi : z \longmapsto z\,\varepsilon$

est alors une *bijection*. Il en résulte que ($\mathbb{Z}$, $+$, $\times$) est isomorphe à (B, $\divideontimes$, $\perp$).

b) Supposons qu'il existe un entier positif non nul tel que $p\,\varepsilon = e$.

Dans ces conditions, les éléments

$$e,\ \varepsilon,\ 2\,\varepsilon,\ \ldots,\ (p-1)\,\varepsilon$$

sont distincts ; on désigne par B cet ensemble d'éléments.

L'application φ de C_p vers B

$$\varphi : [k] \longmapsto k\,\varepsilon \text{ avec } k \in \{0, 1, 2, \ldots, p-1\}$$

est une bijection. C'est un *isomorphisme* de C_p vers (B, $\divideontimes$, $\perp$).

Caractéristique d'un anneau intègre

(A, $\divideontimes$, $\perp$) est un anneau. e est l'élément neutre du groupe (A, $\divideontimes$).
a désigne un élément quelconque de A, et sym a le symétrique de a pour la loi de composition $\divideontimes$.

Posons : $e\ a = e$

$1\ a = a$

$2\ a = a \divideontimes a$

$3\ a = 2\ a \divideontimes a$

...

$n\ a = (n - 1)\ a \divideontimes a$ où n est un naturel.

De même : $(-1)\ a = \text{sym}\ a$

$(-2)\ a = \text{sym}\ a \times \text{sym}\ a$

$(-3)\ a = (-2)\ a \times \text{sym}\ a$

...

$(-n)\ a = (-n + 1)\ a \times \text{sym}\ a$ où n est un naturel.

Notez que la multiplication d'un entier par a est une loi qui ne doit pas être confondue avec la loi de composition $\perp$ dans A.

Nous avons ici une *loi externe*, c'est une application de $\mathbb{Z} \times A$ dans A.

Il est facile de vérifier que quels que soient les entiers p et q et quels que soient les éléments a et b de A :

$$(p + q)\ a = p\ a \divideontimes q\ a$$
$$p\ (q\ a) = (p\ q)\ a$$
$$p\ (a \divideontimes b) = p\ a \divideontimes p\ b$$
$$p\ (a \perp b) = (p\ a) \perp b = a \perp (p\ b)$$

Supposons que $(A, \divideontimes, \perp)$ soit un anneau intègre d'élément unité ε.

e désigne l'élément neutre du groupe $(A, \divideontimes)$.

Quel que soit l'élément x de A et le naturel p

$$p\ x = p\ (\varepsilon \perp x) = (p\ \varepsilon) \perp x$$

Il en résulte que

si $p\ \varepsilon = e$ alors $p\ x = e$

et si $p\ x = e$ et $x \neq e$ alors $p\ \varepsilon = e$

car nous avons ici un anneau intègre.

Ainsi, dans un anneau intègre

ou bien $p\ \varepsilon = e$ et quel que soit x élément de A, $p\ x = e$

ou bien $p\ x \neq e$ pour tout élément x de A, différent de e.

Supposons que, pour tout entier positif p, $p\ \varepsilon \neq e$.

Il en résulte que pour tout élément x de A $(x \neq e)$, $px \neq e$.

On dit alors que l'anneau intègre $(A, \divideontimes, \perp)$ est de *caractéristique nulle*.

C'est le cas de l'anneau des entiers $(\mathbb{Z}, +, \times)$.

Au contraire, s'il existe un naturel non nul (que l'on peut toujours prendre positif) tel que $n\varepsilon = e$, nous savons que dans ce cas, quel que soit l'élément x de A, $nx = e$; en désignant par p le plus petit entier positif, tel que $p\varepsilon = e$, on dit que l'anneau intègre (A, $\divideontimes$, $\perp$) est de *caractéristique* p.

On démontre que p est *premier*.

L'anneau des entiers modulo p (p premier) est de caractéristique p.

A titre d'exercice, le lecteur vérifiera que l'ensemble $\{ e, \varepsilon, a, b \}$ muni des deux lois de composition notées respectivement $\divideontimes$ et $\perp$, données par les tables suivantes :

$\divideontimes$	e	ε	a	b
e	e	ε	a	b
ε	ε	e	b	a
a	a	b	e	ε
b	b	a	ε	e

$\perp$	e	ε	a	b
e	e	e	e	e
ε	e	ε	a	b
a	e	a	b	ε
b	e	b	ε	a

est un anneau intègre de caractéristique 2.

CORPS

Définition

Supposons maintenant que (E, $\divideontimes$, $\perp$) soit un *anneau commutatif unitaire*, tel que tout élément de E, différent de l'élément neutre e pour la loi de composition notée $\divideontimes$ (la première), admette un *symétrique* pour la loi de composition notée $\perp$ (la deuxième).

Dans ce cas, (E, $\divideontimes$, $\perp$) est un *corps*.

En d'autres termes, (E, $\divideontimes$, $\perp$) est un corps si :

1) (E, $\divideontimes$) est un *groupe commutatif* (e désigne l'élément neutre de ce groupe) ;

2) E* désignant l'ensemble E privé de l'élément e, (E*, $\perp$) est un *groupe* (ε désigne l'élément neutre de ce groupe) ;

3) la loi de composition $\perp$ est *distributive* sur la loi de composition $\divideontimes$.

En outre, le corps (E, $\divideontimes$, $\perp$) est *commutatif*, si la loi de composition est commutative.

Exemple

L'ensemble des réels $\mathbb{R}$ muni de l'addition et de la multiplication est un corps.

Propriétés des corps

Un corps est un anneau commutatif unitaire ; il en résulte que toutes les propriétés des anneaux et des groupes sont encore valables pour les corps.
En particulier,

1) L'élément neutre est *unique.*

2) Le symétrique pour la loi de composition notée $\perp$ est *unique* ; si x désigne un élément de E, ce symétrique sera noté x^{-1}.

3) Pour $x \neq e$

$$x \perp y = x \perp z \qquad \text{entraîne} \qquad y = z$$

4) L'équation $a \perp x = b$ possède une solution *unique* : $x = a^{-1} \perp b$ à condition que $a \neq e$.

5) Quel que soit l'élément x de E

$$x \perp e = e \perp x = e$$

6) Quels que soient les éléments x et y de E

$$(x \perp y)^{-1} = y^{-1} \perp x^{-1}$$

7) *Tout corps est un anneau intègre.*
En effet, supposons que, pour deux éléments x et y de E, on ait

$$x \perp y = e \quad \text{avec} \quad y \neq e$$

comme $y \neq e$, y admet un symétrique y^{-1} pour la loi de composition notée $\perp$.
Nous avons alors

$$\begin{aligned} x &= x \perp (y \perp y^{-1}) \\ &= (x \perp y) \perp y^{-1} \\ &= e \perp y^{-1} \\ &= e \end{aligned}$$

Ainsi, $y \neq e$ et $x \perp y = e$ entraîne $x = e$
Il en résulte que (E, $\ast$, $\perp$) est un anneau intègre.

8) L'anneau des entiers modulo p est un *corps* si p est *premier.*

Maurice Glaymann.

L'algèbre linéaire

par Danielle Fèvre

C'est l'étude d'un ensemble muni d'une structure algébrique particulière appelé espace vectoriel. Cette structure a une représentation tellement évidente dans l'espace dans lequel nous vivons que nous en avons fait un domaine a part de l'algèbre : celui qui permet d'aborder d'une façon nouvelle la géométrie euclidienne. Les propriétés géométriques qui nous semblent familières, telles que le parallélisme de deux droites par exemple, se trouvent résolues d'une façon axiomatique et algébrique. Dans ce cas, les démonstrations de ces propriétés géométriques peuvent se passer d'une représentation imagée et se résolvent par des propriétés algébriques uniquement.

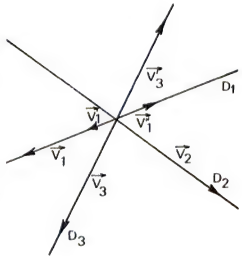
Danielle Fevre.
Née en 1945.
Agrégée de mathématiques, ancienne élève de l'École normale supérieure, maître-assistant à la Faculté des sciences de Metz. Elle fait actuellement des recherches sur l'utilisation des mathématiques en archéologie.

ESPACE VECTORIEL

La notion de vecteur a une représentation assez simple dans le plan. C'est par cette représentation imagée que nous commencerons. Nous aborderons ensuite la définition axiomatique des espaces vectoriels. Choisissons un point O dans le plan ◆. Par ce point, faisons passer des droites. Si nous partons de O, nous pouvons parcourir la droite dans un sens ou dans l'autre. Choisir un sens de référence s'appelle orienter la droite. Ayant orienté la droite et partant toujours de O, nous pouvons parcourir un morceau de la droite dans le sens positif (sens choisi pour l'orientation) ou dans le sens négatif (sens contraire). Ce morceau de droite orientée parcouru depuis O s'appelle *vecteur* (d'origine O). Nous appellerons l'ensemble de tous les vecteurs que nous pouvons définir ainsi depuis O : l'ensemble des vecteurs d'origine O. Nous notons $\vec{V}$ un vecteur de cet ensemble.

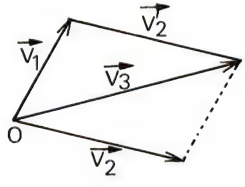
Ayant maintenant à notre disposition cet ensemble de vecteurs, nous allons le munir de deux lois et c'est *la structure particulière que vont lui conférer ces deux lois que nous appellerons espace vectoriel.*

◆ Il est évident que les mots « points », « plan », « droite » n'ont pas été définis ici. Cette partie ne tient lieu que de support imagé de la théorie des espaces vectoriels. Nous verrons par la suite que ces notions peuvent être redéfinies à partir de la construction axiomatique des espaces vectoriels.



1) *Somme de deux vecteurs.*

Ayant deux vecteurs $\vec{V}_1$ et $\vec{V}_2$, nous appelons $\vec{V}_3$ somme de $\vec{V}_1$ et $\vec{V}_2$ le vecteur représenté par la diagonale du parallélogramme de côtés $\vec{V}_1$ et $\vec{V}_2$ et d'origine O.



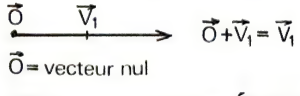
Remarque : pour construire cette diagonale, il revient au même de tracer le vecteur $\vec{V}_1$ et, à l'extrémité de $\vec{V}_1$, de tracer un vecteur $\vec{V}_2'$ parallèle à $\vec{V}_2$, de même sens et de même longueur. L'extrémité de $\vec{V}_2'$ coïncide avec celle de $\vec{V}_3$ ◆.

◆ La somme que nous définissons ici n'est qu'une représentation possible de la somme dans un espace vectoriel. Nous la définirons de façon plus générale par la suite.

Elle est commutative, c'est-à-dire que $\vec{V}_1 + \vec{V}_2 = \vec{V}_2 + \vec{V}_1$.

La somme de $\vec{V}_1$ et $\vec{V}_2$ est égale à la somme de $\vec{V}_2$ et $\vec{V}_1$, quels que soient $\vec{V}_1$ et $\vec{V}_2$. Cela est évident puisque l'ordre dans lequel nous nommons $\vec{V}_1$ et $\vec{V}_2$ ne change pas le dessin du parallélogramme ni le tracé de sa diagonale.

Elle admet un élément neutre $\vec{O}$, c'est-à-dire qu'il existe un vecteur qui, ajouté à tout autre vecteur de l'espace, le laisse inchangé. Ce vecteur est le vecteur nul. Nous appelons ainsi le vecteur réduit à l'origine.



Le parallélogramme est aplati sur $\vec{V}_1$; la diagonale en est $\vec{V}_1$.

Elle est associative. Étudions la somme de 3 vecteurs $\vec{V}_1$, $\vec{V}_2$, $\vec{V}_3$. Par définition, $\vec{V}_1 + \vec{V}_2 + \vec{V}_3$ se calcule en évaluant la somme de $\vec{V}_1$ et $\vec{V}_2$, qui est un vecteur $\vec{W}$, et en lui ajoutant $\vec{V}_3$. Nous obtenons ainsi un vecteur $\vec{S}$. Par définition :

$$\vec{S} = \vec{V}_1 + \vec{V}_2 + \vec{V}_3 = (\vec{V}_1 + \vec{V}_2) + \vec{V}_3 = \vec{W} + \vec{V}_3$$

si $\vec{W} = \vec{V}_1 + \vec{V}_2$.

Supposons que nous calculons la somme de $\vec{V}_2$ et $\vec{V}_3$. Nous obtenons ainsi un vecteur $\vec{U}$. Nous pouvons montrer alors que la somme de $\vec{V}_1$ et de $\vec{U}$ est encore le vecteur $\vec{S}$, ce qui s'écrit :

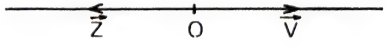
$$\vec{S} = \vec{V}_1 + (\vec{V}_2 + \vec{V}_3) = \vec{V}_1 + \vec{U} \quad (\text{si } \vec{U} = \vec{V}_2 + \vec{V}_3)$$

D'où la propriété :

$$\boxed{(\vec{V_1} + \vec{V_2}) + \vec{V_3} = \vec{V_1} + (\vec{V_2} + \vec{V_3})}$$

quels que soient $\vec{V_1}$, $\vec{V_2}$ et $\vec{V_3}$.

Chaque vecteur admet un inverse, c'est-à-dire que, quel que soit le vecteur $\vec{V}$ considéré, nous pouvons toujours trouver un vecteur $\vec{Z}$ tel que $\vec{V} + \vec{Z} = \vec{Z} + \vec{V} = \vec{O}$. En effet, il suffit pour cela de prendre le vecteur $\vec{Z}$ porté par la même droite que $\vec{V}$, de même longueur que $\vec{V}$ et de sens contraire. Dans ce cas, le parallélogramme construit sur $\vec{V}$ et $\vec{Z}$ est aplati, ayant deux sommets confondus en O. La diagonale est bien le vecteur nul :



L'addition que nous venons de définir dans l'ensemble des vecteurs considérés a les propriétés suivantes : elle est interne, elle est associative, elle admet un élément neutre, chaque élément admet un inverse. Cela confère à l'ensemble des vecteurs une structure de groupe. De plus, la loi est commutative. L'espace a donc une structure de groupe commutatif ◆.

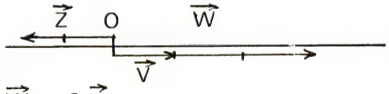
◆ Ces propriétés sont celles qui définissent un groupe. Voir l'article « l'Algèbre », pages 28 à 45. Groupe commutatif y est également défini.

Nous retiendrons ces cinq propriétés et elles nous serviront dans la définition axiomatique des espaces vectoriels : *un espace vectoriel est muni d'une première loi interne qui lui donne une structure de groupe commutatif.*

2) *Produit d'un vecteur par un nombre réel.*

Considérons un vecteur $\vec{V}$ de notre ensemble. Le multiplier par 3 consistera à le transformer en un vecteur $\vec{W}$ de même support que lui, de même direction et trois fois plus long. Le multiplier par — 1,6 consistera à le transformer en un vecteur $\vec{Z}$ de même support que lui, de sens contraire, et dont la longueur vaudra + 1,6 fois celle de $\vec{V}$. Multiplier $\vec{V}$ par un nombre réel quelconque α sera le transformer en un vecteur $\vec{T}$, porté par le même support que $\vec{V}$, de même sens si le nombre est positif, de sens contraire s'il est négatif, et dont la longueur vaudra $|\alpha|$ ◆ fois celle de $\vec{V}$.

◆ Cette notation $|\alpha|$ indique qu'il s'agit de la valeur absolue de α (+ α si α est positif, — α si α est négatif).



$\vec{W} = 3\,\vec{V}$

$\vec{Z} = -\,1\,6\,\vec{V}$

S'il existe un nombre réel tel que $\vec{W} = \alpha\,\vec{V}$, alors nous disons que $\vec{W}$ est proportionnel ou colinéaire à $\vec{V}$.

Cette opération a-t-elle des propriétés particulières? La première remarque, qui semble une évidence et qui pourtant est une propriété due à celle des nombres réels, est la suivante :

$$1 \cdot \vec{V} = \vec{V}$$

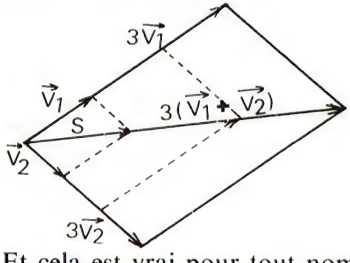
En effet, si nous multiplions un vecteur $\vec{V}$ par 1, nous le transformons en un vecteur porté par le même support, de même direction (puisque 1 est positif) et de même longueur. D'où le résultat : $1 \cdot \vec{V} = \vec{V}$.

L'ensemble des nombres réels $\mathbb{R}$ est déjà muni de deux lois ◆, une addition et une multiplication, et l'ensemble des vecteurs est muni d'une loi interne : l'addition vectorielle. Nous pouvons nous demander si la multiplication d'un vecteur par un nombre réel a des propriétés particulières relativement à ces trois lois. Commençons par les additions.

◆ Voir l'article « l'Algèbre », pages 17 à 27, et l'article « le Nombre », page...

La multiplication d'un vecteur par un nombre réel est distributive par rapport à l'addition des nombres réels, d'une part, et par rapport à l'addition des vecteurs, d'autre part.

Supposons que nous ayons à calculer $3\,(\vec{V_1} + \vec{V_2})$. Nous calculons d'abord $\vec{V_1} + \vec{V_2}$, qui nous donne un vecteur $\vec{S}$, et nous le multiplions par 3. Nous obtenons ainsi un vecteur $\vec{T}$. Or nous aurions obtenu le même résultat si nous avions d'abord calculé $3\,\vec{V_1}$ puis $3\,\vec{V_2}$ et additionné ces 2 vecteurs.



En effet, $\vec{T}$ peut être considéré comme la diagonale du parallélogramme de côtés $3\,\vec{V_1}$ et $3\,\vec{V_2}$.

Et cela est vrai pour tout nombre réel. Nous avons donc :

$$\forall\, \alpha \in \mathbb{R} \quad \forall\, \vec{V_1} \text{ et } \vec{V_2} \in E \quad \alpha(\vec{V_1} + \vec{V_2}) = \alpha\,\vec{V_1} + \alpha\,\vec{V_2}\ ◆.$$

◆ Lire : quel que soit x appartenant à l'ensemble des nombres réels, quels que soient les vecteurs V_1 et V_2 appartenant à l'ensemble E, etc.

Nous disons que la multiplication par un nombre réel est distributive par rapport à l'addition des vecteurs.

Supposons que nous ayons à calculer $(3 + 4)\,\vec{V}$. C'est un vecteur de même support que $\vec{V}$, de même direction puisque $(3 + 4)$ est positif et dont la longueur vaut $3 + 4 = 7$ fois celle de $\vec{V}$. Nous pouvons dire aussi que c'est la somme des vecteurs $3\,\vec{V} + 4\,\vec{V}$. Cette somme sera bien le vecteur $7\,\vec{V}$. Nous avons la propriété :

$$(3 + 4)\,\vec{V} = 3\,\vec{V} + 4\,\vec{V}.$$

Et, plus généralement :

$$\forall\ \alpha_1 \text{ et } \alpha_2 \in \mathbb{R} \quad \forall\ \vec{V} \in E \quad (\alpha_1 + \alpha_2)\,\vec{V} = \alpha_1\,\vec{V} + \alpha_2\,\vec{V}.$$

Nous disons que la multiplication d'un vecteur par tout nombre réel est distributive par rapport à l'addition des nombres réels.

Dernière propriété : associativité des nombres réels pour la multiplication d'un vecteur par un nombre réel.

Regardons ce que devient le vecteur $\vec{V_1}$ quand nous le multiplions par le nombre 3×4. Nous avons alors :

$$(3 \times 4)\,\vec{V_1} = 12\,\vec{V_1}.$$

Nous pouvons dire aussi que le vecteur $12\,\vec{V_1}$ est le vecteur $4\,\vec{V_1}$ multiplié par 3, car les deux vecteurs $12\,\vec{V_1}$ et $3.(4\,\vec{V_1})$ ont bien même support, même direction et même longueur. Nous avons donc la propriété : $(3 \times 4)\,\vec{V_1} = 3.(4\,\vec{V_1})$.

Et, plus généralement :

$$\forall\ \alpha \text{ et } \beta \in \mathbb{R} \quad \forall\ \vec{V} \in E \quad (\alpha \times \beta).\,\vec{V} = \alpha.(\beta.\,\vec{V}).$$

Nous disons qu'il y a associativité des nombres réels pour la multiplication.

La multiplication par un nombre réel, que nous venons de définir, a les 4 propriétés suivantes :
— elle est distributive par rapport à l'addition des vecteurs ;
— elle est distributive par rapport à l'addition des nombres réels ;
— elle est associative pour la multiplication des nombres réels ;
— l'élément neutre pour la multiplication des nombres réels est aussi élément neutre pour la multiplication des vecteurs par un nombre réel.

Ces propriétés aussi nous serviront dans la définition axiomatique des espaces vectoriels.

Définition axiomatique

Récapitulons les propriétés de l'espace vectoriel. Par définition, un espace vectoriel est un ensemble muni de deux opérations.

— *La première opération est interne :* c'est-à-dire qu'à deux vecteurs $\vec{V_1}$ et $\vec{V_2}$ elle permet d'en associer un troisième. Elle est associative et commutative, elle admet un élément neutre et chaque élément admet un inverse.

— *La deuxième opération est externe :* à un nombre et à un vecteur donnés, on sait associer un vecteur. Mais ce nombre n'est pas nécessairement un nombre réel. La seule hypothèse à faire sur l'ensemble de ces éléments étrangers aux vecteurs est qu'il ait une structure de corps ◆. On le note souvent K ; on l'appelle le *corps des scalaires* (dans l'exemple précédent, ces scalaires étaient des nombres réels, comme c'est souvent le cas). Cette opération externe a les propriétés suivantes : elle admet un élément neutre dans les scalaires, elle est distributive par rapport à l'addition des vecteurs et par rapport à l'addition des scalaires, elle est associative pour la multiplication des scalaires entre eux par un vecteur.

◆ Voir l'article « l'Algèbre », page 53.

— Il existe encore une propriété remarquable des espaces vectoriels, qui est une conséquence des propriétés précédentes. Si $\vec{O}$ est le vecteur nul, élément neutre de l'addition vectorielle, alors l'égalité $\alpha \vec{V} = \vec{O}$ entraîne que soit $\alpha = 0$, soit $\vec{V} = \vec{O}$. Réciproquement, si $\alpha = 0$ ou $\vec{V} = \vec{O}$, alors $\alpha \vec{V} = \vec{O}$.

Remarque très importante

On a tendance à considérer que les vecteurs sont des segments de droite orientés. L'exemple que nous venons de traiter aurait tendance à confirmer cette idée. Cette représentation de l'espace vectoriel n'en est qu'un cas particulier. Il existe beaucoup d'autres représentations d'espaces vectoriels. Nous allons en donner quelques exemples, où la représentation des vecteurs n'est plus un segment de droite orienté : l'espace vectoriel peut être un espace de fonctions, de couples ou de triplets, de nombres réels, etc.

Exemples d'espaces vectoriels

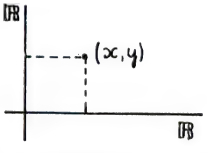
Pour bien comprendre ces exemples, il est recommandé de lire, d'une part, l'article « l'Algèbre », pour les notions de groupe,

d'anneau, de corps, et l'article « l'Analyse », pour la définition d'une fonction, de sa dérivée et de sa primitive.
La lecture de ces exemples n'est pas indispensable pour comprendre la suite de l'article, mais il est conseillé d'y revenir pour bien saisir la notion d'espace vectoriel.

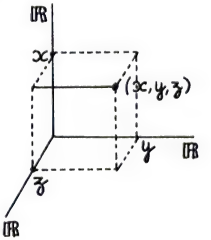
Premier exemple

Considérons l'ensemble des triplets de nombres réels, noté $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ ◆. C'est l'ensemble des éléments de la forme (x, y, z), où x et y et z appartiennent à l'ensemble $\mathbb{R}$ des nombres réels. Nous définissons ainsi l'addition dans $\mathbb{R}^3$.

◆ Le couple s'écrit (x, y). Une représentation imagée possible est :



Le triplet s'écrit (x, y, z).



$$\overbrace{(x,\ y,\ z)}^{1^{er}\ \text{triplet}} + \overbrace{(x',\ y',\ z')}^{2^{e}\ \text{triplet}} = (x + x',\ y + y',\ z + z').$$

Loi interne
A cause des propriétés de $\mathbb{R}$, cette addition dans $\mathbb{R}^3$ est bien associative, commutative et admet un élément neutre $(0, 0, 0)$ et chaque élément admet un inverse $[(x, y, z) + (-x, -y, -z) = (0, 0, 0)$; l'inverse de (x, y, z) est $(-x, -y, -z)]$. $\mathbb{R}^3$ a une structure de groupe commutatif pour cette addition.

Loi externe
Nous pouvons définir la multiplication d'un élément de $\mathbb{R}^3$ par un nombre réel α ; nous posons :

$$\alpha(x, y, z) = (\alpha x, \alpha y, \alpha z).$$

C'est bien une opération externe de $\mathbb{R}^3$. Grâce aux propriétés de $\mathbb{R}$, elle jouit des propriétés : distributivité de la multiplication d'un triplet par un scalaire par rapport à l'addition des scalaires, par rapport à l'addition des triplets, associativité de la multiplication des scalaires et

$$1 \,.(x,\ y,\ z) = (x,\ y,\ z)$$

quel que soit le triplet. $\mathbb{R}^3$ est muni d'une structure d'espace vectoriel.

Deuxième exemple

On peut montrer également que l'ensemble F des fonctions d'une variable réelle à valeur réelle ◆ peut être muni d'une structure d'espace vectoriel.

◆ Voir l'introduction de l'article « l'Analyse ».

En effet, nous savons définir la somme de 2 applications f et g définies sur $\mathbb{R}$ et à valeurs réelles dans $\mathbb{R}$; c'est l'application h qui, à tout x appartenant à $\mathbb{R}$, associe la valeur :

$$h(x) = f(x) + g(x)$$
$$f + g = h : x \longmapsto f(x) + g(x)$$

Cette addition donne à F une structure de groupe commutatif. En effet, on peut montrer qu'elle est associative, commutative, qu'il

existe un élément neutre (c'est l'application nulle, application qui à tout x réel associe le nombre réel 0) et que chaque fonction f admet une fonction opposée $-f$ $(f + (-f) = 0$.

Nous savons également définir la multiplication d'une telle fonction f par un nombre réel α ; c'est l'application qui à tout x appartenant à $\mathbb{R}$ associe le nombre $\alpha . f(x)$:

$$\alpha . f : x \longmapsto \alpha f(x)$$

Cette deuxième loi vérifie bien la propriété de définition de l'espace vectoriel : distributivité de la multiplication d'une fonction par un nombre réel par rapport à l'addition des fonctions, à l'addition des nombres réels, associativité de la multiplication des nombres réels. Et, si 1 désigne l'élément neutre de la multiplication des nombres réels :

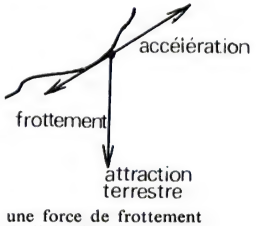
$$1 . f = f$$

Cet ensemble F muni de 2 lois : addition et multiplication, par un scalaire réel, a bien une structure d'espace vectoriel sur $\mathbb{R}$.

APPLICATIONS

Cette notion d'espace vectoriel intervient dans des domaines très divers, en dynamique en particulier ◆.

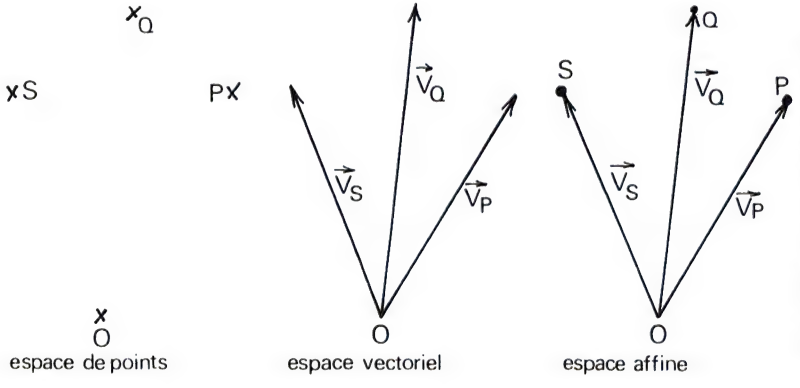
◆ Supposons que nous voulions étudier le système de forces auxquelles est soumise une petite balle de plomb lancée en l'air, qu'on peut accélérer en cours de trajet. Elle sera, à un moment donné, soumise à 3 forces :



une force de frottement de l'air s'opposant à son mouvement, une force d'attraction terrestre l'attirant vers le bas et la force d'accélération qu'on peut lui donner quand on veut. Suivant la convention qu'une force peut être représentée par un vecteur, nous avons alors la représentation suivante : On montre alors que l'ensemble de ces 3 forces a le même effet qu'une seule force telle que le vecteur la représentant, soit la somme vectorielle des 3 autres.

Espace affine

Une autre utilisation courante des espaces vectoriels est l'emploi des vecteurs pour situer les points de l'espace ordinaire. On choisit une origine O dans l'espace, et tout point M pourra être représenté par le vecteur d'origine O et d'extrémité M.
S'il existe entre les points de l'espace des relations se ramenant à des relations simples entre les vecteurs les représentant, il est recommandé alors d'utiliser les éléments de calcul vectoriel que nous venons de définir. Il faut bien noter que l'ensemble des points et



l'espace vectoriel qui permet de les décrire sont deux êtres mathématiques distincts. L'espace des points sur lequel il est possible de travailler grâce à la structure d'espace vectoriel sous-jacente s'appelle alors *espace affine*.
Pour avoir une idée des possibilités qui nous sont offertes dans un espace affine, nous allons donner, dans le langage du calcul vectoriel, quelques définitions bien connues de la géométrie d'Euclide ◆.

◆ Nous utilisons les propriétés de la géométrie traditionnelle sans les démontrer. En revanche, les définitions vectorielles que nous donnons permettent une définition rigoureuse.

droite affine

droite vectorielle

Nous appelons droite affine passant par O *et de direction* $\vec{U}$ *(où* $\vec{U}$ *est un vecteur d'origine* O*) les extrémités (les points) de l'ensemble des vecteurs de la forme* $\alpha \vec{U}$ *où* α *décrit l'ensemble des nombres réels.*
L'ensemble des vecteurs de la forme $\alpha \vec{U}$ *s'appelle droite vectorielle de direction* $\vec{U}$.

D'après la définition même de la multiplication d'un vecteur par un nombre réel, tous ces points se trouvent sur la droite portant $\vec{U}$. Et, réciproquement, tout point M de cette droite définit un vecteur colinéaire ◆ à $\vec{U}$; dans ce cas, nous saurons trouver un nombre réel tel que $\overrightarrow{OM} = \beta \vec{U}$ (β sera positif ou négatif suivant que $\overrightarrow{OM}$ et $\vec{U}$ sont de même sens ou de sens contraire et sa valeur est définie par le rapport de la longueur de $\overrightarrow{OM}$ à celle de $\vec{U}$).

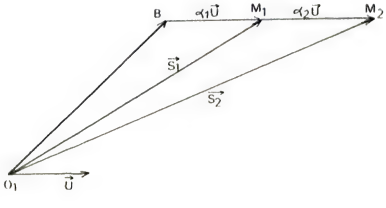
◆ Voir définition, ci-dessus, page 58.

Nous pouvons aussi définir une droite affine parallèle à un vecteur donné $\vec{U}$ et passant par un point B quelconque. Soit O_1 l'origine du vecteur $\vec{U}$ donné. *Nous appelons droite affine parallèle à* $\vec{U}$ *passant par* B *l'ensemble des points* M *extrémités des vecteurs ainsi définis :*

$$\overrightarrow{O_1M} = \vec{S} = \overrightarrow{O_1B} + \alpha \vec{U}$$

où α *décrit l'ensemble des nombres réels* ◆.

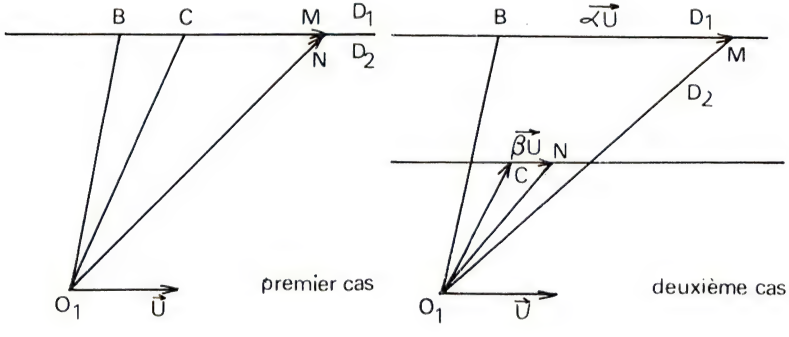
◆ Voir plus haut l'addition des vecteurs.



Dans ce cas, nous pouvons dire que $\overrightarrow{BM}$ est parallèle à $\vec{U}$.

Nous pouvons démontrer que, sous ces définitions, *deux droites parallèles à un même vecteur* $\vec{U}$ *et passant respectivement par 2 points distincts* B *et* C *n'ont aucun point commun ou sont confondues.*

Cela dépend de la position de BC par rapport à $\vec{U}$.



1er *cas* : BC *parallèle à* $\vec{U}$.

Si BC est parallèle à $\vec{U}$, nous avons la relation :

$$\boxed{\overrightarrow{O_1C} = \overrightarrow{O_1B} + \gamma\vec{U}}$$

d'après la définition précédente. Les points M de la droite D_1 passant par B sont définis par :

$$\overrightarrow{O_1M} = \overrightarrow{O_1B} + \alpha\vec{U}$$

et les points N de la droite D_2 passant par C par :

$$\overrightarrow{O_1N} = \overrightarrow{O_1C} + \beta\vec{U}$$

Or $\overrightarrow{O_1C} = \overrightarrow{O_1B} + \gamma\vec{U}$. Aussi :

$$\overrightarrow{O_1N} = (\overrightarrow{O_1B} + \gamma\vec{U}) + \beta\vec{U} = \overrightarrow{O_1B} + (\gamma\vec{U} + \beta\vec{U})$$
$$= \overrightarrow{O_1B} + (\gamma + \beta)\vec{U} \text{ (définition de M)}$$

Cela montre que N est sur la droite D_1, et tout point de la droite D_2 se trouve sur D_1.

De la même façon, nous pouvons montrer que tout point de D_1 se trouve sur D_2.

Dans ce cas, D_1 *et* D_2 *sont confondues.*

2e *cas* : BC *n'est pas parallèle à* $\vec{U}$.

Nous ne pouvons pas trouver M et N tels que $\overrightarrow{O_1M} = \overrightarrow{O_1N}$.

En effet, pour cela, il faudrait avoir :

$$\overrightarrow{O_1B} + \alpha\vec{U} = \overrightarrow{O_1C} + \beta\vec{U}$$

soit : $\overrightarrow{O_1B} - \overrightarrow{O_1C} = \beta\vec{U} - \alpha\vec{U} = (\beta - \alpha)\vec{U}$

On peut montrer par le calcul vectoriel que cette égalité n'est pas réalisable si BC n'est pas parallèle à $\vec{U}$.

Dans ce cas, D_1 *et* D_2 *n'ont aucun point commun.*

BASE D'UN ESPACE VECTORIEL

Nous venons de décrire un ensemble très précis : l'ensemble des vecteurs d'origine O, qui contient un nombre impressionnant de vecteurs. Nous venons de le munir de 2 lois : l'addition des vecteurs et la multiplication d'un vecteur par un nombre réel ◆, c'est-à-dire que nous savons ce que représentent les expressions suivantes :

◆ Nous reprenons le cas d'un espace vectoriel défini sur l'ensemble des nombres réels. Pour tout autre corps de scalaires, les propriétés seraient les mêmes.

$$\alpha\vec{V_1}, \quad \vec{V_1} + \vec{V_2}, \quad \alpha\vec{V_1} + \vec{V_2}, \quad \alpha\vec{V_1} + \beta\vec{V_2}$$

et, si nous en considérons davantage :

$$\alpha_1\vec{V_1} + \alpha_2\vec{V_2} + \alpha_3\vec{V_3} + \alpha_4\vec{V_4} + \alpha_5\vec{V_5}$$

par exemple. Cela nous permet une combinaison extrêmement complexe de vecteurs et de nombres réels.

Toute combinaison obtenue en utilisant ces deux opérations s'appelle combinaison linéaire des vecteurs.

Maintenant que nous sommes plus familiarisés avec la notion d'espace vectoriel, nous adopterons les conventions suivantes : une lettre majuscule sans flèche représente un vecteur, une lettre minuscule latine ou grecque un scalaire.

Quelle que soit la complexité de l'expression, telle celle-ci :

$$(a + b)(V_1 + cV_2) + c(a \,.\, r)\,V_1 + [m(V_1 + V_2 + V_3 + V_4) + n(V_6 + V_1)]\,r$$

elle peut toujours se ramener, en utilisant les propriétés de distributivité et d'associativité, à une expression de la forme :

$$a_1V_1 + a_2V_2 + a_3V_3 + a_4V_4 + \ldots + a_nV_n$$

(V_1, V_2, V_n) correspondant aux n vecteurs différents rencontrés dans l'expression donnée. Pour s'en convaincre, il suffit d'essayer sur un exemple.

SYSTÈMES GÉNÉRATEURS

Devant le grand nombre de vecteurs contenus dans un espace vectoriel, nous pouvons nous demander s'il n'est pas possible d'exhiber

un petit nombre d'entre eux tel que tout vecteur de l'espace vectoriel puisse s'exprimer comme une combinaison linéaire de ceux-ci. Dans ce cas, seule la connaissance de ce petit nombre de vecteurs nous permettrait de connaître tout l'espace. Nous ne traiterons pas ici de la résolution de ce problème. Nous admettrons qu'en général cela est possible, et nous le verrons effectivement sur des exemples particulièrement simples.
Soit V un vecteur de l'espace vectoriel E considéré. Supposons que V soit une combinaison linéaire de 4 vecteurs V_1, V_2, V_3, V_4, par exemple, c'est-à-dire qu'il existe 4 nombres réels a_1, a_2, a_3, a_4 tels que

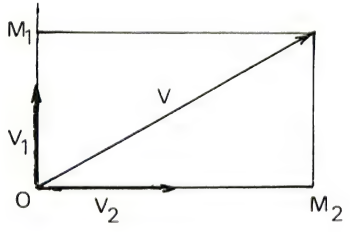
$$V = a_1V_1 + a_2V_2 + a_3V_3 + a_4V_4$$

Dans ce cas, nous disons que :
— V est combinaison linéaire de (V_1, V_2, V_3, V_4) ;
— le système (V_1, V_2, V_3, V_4) engendre V ;
— le système (V_1, V_2, V_3, V_4) est un système générateur de V.

Et, si tout vecteur de E peut s'exprimer comme combinaison linéaire des 4 vecteurs V_1, V_2, V_3, V_4, nous disons alors que :
— le système (V_1, V_2, V_3, V_4) engendre E ;
— le système (V_1, V_2, V_3, V_4) est un système générateur de E.

Nous avons donné ces définitions pour un système formé de 4 vecteurs V_1, V_2, V_3, V_4. Il est évident que les définitions n'ont rien à voir avec ce nombre. Il existe des systèmes générateurs de 1, 2, 3, 4, 5, 6, etc., vecteurs.

Exemple :
Revenons à notre exemple de l'ensemble des vecteurs d'origine O situés dans un plan. Prenons 2 vecteurs perpendiculaires V_1 et V_2. Alors, tout vecteur d'origine O peut être considéré comme la somme de α_1V_1 et α_2V_2, où α_1 et α_2 sont choisis correctement.



Il suffit pour cela, de l'extrémité de V, de tracer la parallèle à V_1 et la parallèle à V_2. Elles coupent respectivement le support de V_2 et V_1 aux points M_2 et M_1. Nous savons évaluer α_1 tel que :

$$OM_1 = \alpha_1V_1$$

et α_2 tel que :

$$OM_2 = \alpha_2V_2$$

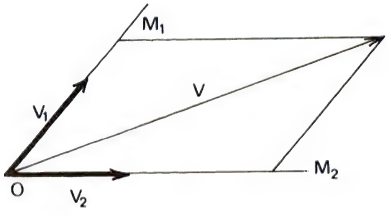
Dans ce cas :

$$V = OM_1 + OM_2 = \alpha_1V_1 + \alpha_2V_2$$

Nous venons de déterminer α_1 et α_2 tel que

$$V = \alpha_1V_1 + \alpha_2V_2$$

et cela quel que soit V. Le système (V_1, V_2) engendre l'ensemble E des vecteurs du plan d'origine O.



Nous pouvons aussi prendre deux vecteurs particuliers V_1 et V_2 non perpendiculaires, mais non portés par la même droite. Suivant le même procédé de construction, nous arriverons à déterminer β_1 et β_2 pour que :

$$V = \beta_1 V_1 + \beta_2 V_2$$

Nous venons de trouver un deuxième système générateur de même espace vectoriel E.

SYSTÈME LIBRE ET SYSTÈME LIÉ

Une deuxième notion très intéressante est celle de l'indépendance des vecteurs. Supposons que nous soyons arrivés à montrer la relation suivante :

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3 + \alpha_4 V_4$$

Nous pourrions nous poser alors la question suivante : est-il possible de trouver une deuxième série de nombres réels β_1, β_2, β_3, β_4 distincts des précédents tels que :

$$V = \beta_1 V_1 + \beta_2 V_2 + \beta_3 V_3 + \beta_4 V_4$$

Faisons la différence entre ces deux expressions :

$$\begin{aligned} V - V &= (\alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3 + \alpha_4 V_4) \\ &\quad - (\beta_1 V_1 + \beta_2 V_2 + \beta_3 V_3 + \beta_4 V_4) \\ O &= (\alpha_1 - \beta_1) V_1 + (\alpha_2 - \beta_2) V_2 + (\alpha_3 - \beta_3) V_3 \\ &\quad + (\alpha_4 - \beta_4) V_4. \end{aligned}$$

La question que nous nous posions revient à la suivante : existe-t-il des nombres réels $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ non tous nuls tels que :

$$O = \gamma_1 V_1 + \gamma_2 V_2 + \gamma_3 V_3 + \gamma_4 V_4$$

avec $\gamma_1 = \alpha_1 - \beta_1$, $\gamma_2 = \alpha_2 - \beta_2$, etc.

Cherchons à résoudre l'équation suivante :

$$O = \gamma_1 V_1 + \gamma_2 V_2 + \gamma_3 V_3 + \gamma_4 V_4$$

où V_1, V_2, V_3, V_4 sont donnés.

Il y a une solution évidente, que nous appellerons la solution *triviale* : c'est $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$. En effet :

$$O = 0V_1 + 0V_2 + 0V_3 + 0V_4$$

A part cela, deux cas peuvent se produire :

Ou bien il existe 4 coefficients $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ *non tous nuls solution de l'équation*

$$\gamma_1 V_1 + \gamma_2 V_2 + \gamma_3 V_3 + \gamma_4 V_4 = O$$

c'est-à-dire qu'il existe une combinaison linéaire ◆ entre ces vecteurs, permettant d'obtenir le vecteur nul. Dans ce cas, nous disons que *le système est lié* et un vecteur V peut se décomposer de plusieurs façons différentes sur le système, car, si

◆ Voir ci-dessus.

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3 + \alpha_4 V_4$$

nous pouvons aussi dire que :

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3 + \alpha_4 V_4 + \underbrace{\gamma_1 V_1 + \gamma_2 V_2 + \gamma_3 V_3 + \gamma_4 V_4}_{O}$$

$$= (\alpha_1 + \gamma_1) V_1 + (\alpha_2 + \gamma_2) V_2 + (\alpha_3 + \gamma_3) V_3 + (\alpha_4 + \gamma_4) V_4$$

Ou bien seuls les coefficients tous nuls sont solution :

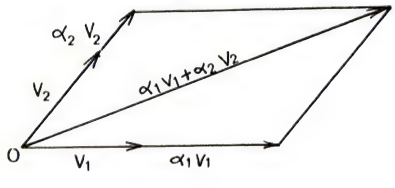
$$0V_1 + 0V_2 + 0V_3 + 0V_4 = O$$

Dans ce cas, *le système est libre.*

Exemples

Reprenons notre exemple ; les 2 systèmes de vecteurs que nous avons étudiés sont libres (voir les 2 figures précédentes). En effet, il sera impossible de trouver α_1 et α_2 non nuls tels que

$$\alpha_1 V_1 + \alpha_2 V_2 = O$$

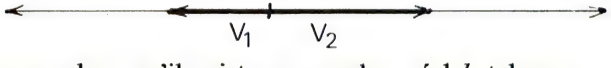


$\alpha_1 V_1$ sera toujours porté par la droite portant V_1, $\alpha_2 V_2$ par celle portant V_2 et nous n'arriverons pas à obtenir une somme nulle, à moins que

$$\alpha_1 V_1 = O \text{ et } \alpha_2 V_2 = O,$$

et cela ne peut se produire que si $\alpha_1 = \alpha_2 = 0$.

Considérons maintenant le système (V_1, V_2), où V_1 et V_2 sont portés par la même droite :



Nous savons alors qu'il existe un nombre réel k tel que :

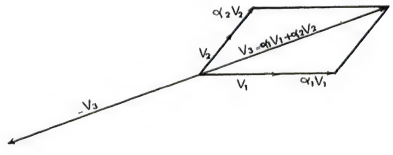
$$V_2 = kV_1$$

Nous avons la relation $-kV_1 + V_2 = O$. Il existe une relation distincte de la relation triviale $0V_1 + 0V_2 = 0$ permettant d'engendrer O.

Le système est lié.

Ou bien prenons trois vecteurs d'origine O, V_1, V_2, V_3, portés par trois droites distinctes passant par O. Nous savons qu'il existe deux nombres α_1 et α_2 tels que :

$$V_3 = \alpha_1 V_1 + \alpha_2 V_2$$



(voir l'exemple traité pour le système générateur). Nous avons donc la relation :

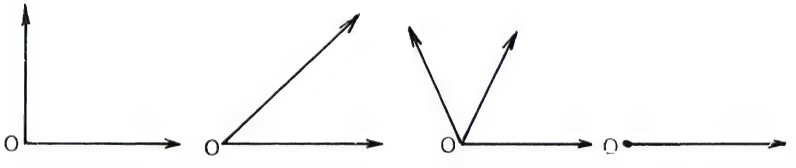
$$\alpha_1 V_1 + \alpha_2 V_2 + (-V_3) = O.$$

Le système est lié.

La décomposition d'un vecteur sur un système libre est plus intéressante que sur un système lié. En effet, dans ce cas, nous sommes assurés de l'unicité de la décomposition possible, nous sommes sûrs par avance que c'est la seule possible.

BASE

Nous voyons apparaître l'importance que prendrait un système de vecteurs qui soit à la fois libre et générateur de l'espace vectoriel E. Un tel système s'appelle une base. Nous avons déjà rencontré des bases, des systèmes générateurs qui ne sont pas des bases (pas libres) et des systèmes libres qui ne sont pas des bases (pas générateurs). Ainsi, si E est l'ensemble des vecteurs d'origine O situés dans un plan donné, le système formé de deux vecteurs non colinéaires d'origine O est une base (nous avons montré précédemment qu'il est générateur et libre) ; le système formé de trois vecteurs non colinéaires deux à deux, d'origine O, est un système générateur et non libre ; le système formé d'un seul vecteur d'origine O est libre et non générateur.



base de E | base de E | système générateur lié de E | système libre non générateur de E

La connaissance des vecteurs d'une base d'un espace vectoriel E nous permet de décrire tout vecteur de E (système générateur) et la décomposition obtenue sur cette base est unique.

Ayant déterminé une base B de l'espace vectoriel E formée de n vecteurs $(V_1, V_2, \ldots, V_n)$, pour tout vecteur V de E nous saurons trouver n scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ (particuliers à chaque vecteur V) tels que V puisse s'écrire :

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n.$$

Les n scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ s'appellent les composantes du vecteur V sur la base B.

Dimensions

Il existe des espaces vectoriels où les bases sont formées d'un seul vecteur : la droite vectorielle, par exemple.

D'autres ont une base formée de deux vecteurs. Nous venons de le voir pour E, espace vectoriel des vecteurs d'origine O situés dans un plan.

D'autres encore ont une base formée de trois vecteurs. Prenons l'ensemble $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ des triplets de nombres réels rencontrés précédemment ◆. On peut montrer facilement que les trois triplets

◆ Voir page 61.

$$E_1 = (1, 0, 0) \qquad E_2 = (0, 1, 0) \qquad E_3 = (0, 0, 1)$$

forment un système libre (pour avoir $\alpha_1 E_1 + \alpha_2 E_2 + \alpha_3 E_3 = (0, 0, 0)$, il faut nécessairement avoir $\alpha_1 = \alpha_2 = \alpha_3 = 0$).

Ce système est générateur, car tout triplet (a, b, c) peut s'écrire :

$$(a, b, c) = (a, 0, 0) + (0, b, 0) + (0, 0, c) = aE_1 + bE_2 + cE_3.$$

Le système (E_1, E_2, E_3) est une base de $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$. D'autres espaces vectoriels admettent une base formée de quatre vecteurs ou d'un nombre n de vecteurs. Il en existe même où il n'est pas possible de trouver une base ayant un nombre fini de vecteurs. L'ensemble $\mathcal{F}$ de toutes les fonctions définies sur l'ensemble des nombres réels à valeurs réelles est ainsi.

Les espaces vectoriels qui admettent une base formée d'un nombre fini de vecteurs jouissent de propriétés particulières que nous allons énoncer ici.

Définition. Soit E un espace vectoriel défini sur un corps K ◆ ; on dit que E est de dimension finie sur K s'il existe une partie Q finie et génératrice de E.

◆ Voir ci-dessus, page 55.

Théorème
Si E (espace vectoriel) est de dimension finie, alors on peut en déduire que :
1) E admet une base B_0.
2) Cette base est finie.
3) Toute autre base de E admet le même nombre de vecteurs que B_0.

Définition. Si B_0 contient n vecteurs, nous disons que E est de dimension n.

Dans un espace vectoriel E, nous avons les propriétés suivantes :

E est de dimension finie $n \Leftrightarrow$ E admet au moins une base formée de n vecteurs et toute autre base de E contient n vecteurs.

Cela entraîne les propriétés équivalentes suivantes :

Si E est de dimension n :

P est une partie libre et contient n vecteurs $\Leftrightarrow$ P est une partie génératrice et contient n vecteurs $\Leftrightarrow$ P est une partie libre et génératrice (base)

Un espace vectoriel de dimension 1 s'appelle une droite vectorielle. Nous l'avons déjà rencontré précédemment. Un espace vectoriel de dimension 2 s'appelle un plan vectoriel. Nous en avons déjà utilisé un très souvent : l'ensemble des vecteurs d'un plan (au sens traditionnel) d'origine O.

Nous voyons que nos définitions algébriques nous permettent une nouvelle définition de la notion de droite et de plan. De la même façon que nous avons défini la droite affine à partir de la droite vectorielle, nous pouvons définir le plan affine à partir du plan vectoriel.

L'espace ordinaire peut être représenté comme un espace vectoriel de dimension 3 (si nous choisissons une origine O fixe).
Einstein, lui, s'est placé dans un espace vectoriel de dimension 4 (trois dimensions pour l'espace ordinaire, une pour le temps) pour résoudre les problèmes de dynamique relativiste.
Il existe de même des espaces de dimension infinie.

ÉCRITURE MATRICIELLE DES VECTEURS

Considérons un espace vectoriel E dont l'ensemble des scalaires est l'ensemble des nombres réels et qui admet une base B formée de n vecteurs : V_1, V_2, ..., V_n. Tous les calculs suivants s'effectueront relativement à cette base B.
Nous venons de voir que, dans ce cas, tout vecteur V de E peut se décomposer sur cette base B, c'est-à-dire que, quel que soit V, nous pouvons déterminer n nombres réels α_1, α_2, ..., α_n tels que :

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n.$$

Nous savons de plus que ces nombres $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ forment un système unique pour chaque vecteur V, c'est-à-dire qu'il n'est pas possible de trouver un second système $(\beta_1, \beta_2, \ldots, \beta_n)$ distinct du précédent et vérifiant :

$$V = \beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_n V_n$$

Nous avons déjà vu que ce système $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ s'appelle système des composantes de V sur la base B ◆.

◆ Voir page 65.

Réciproquement, nous nous donnons un système $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ de n nombres réels ; nous pouvons considérer l'expression

$$\gamma_1 V_1 + \gamma_2 V_2 + \ldots + \gamma_n V_n,$$

qui est nécessairement un vecteur de E. Donc, à un système de n nombres réels $(\gamma_1, \gamma_2, \ldots, \gamma_n)$, nous savons associer un vecteur de E. Un vecteur V de E peut être représenté par le système $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ correspondant, que nous écrivons sous forme d'une matrice à une colonne et n lignes :

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$$ Matrice des composantes de V sur la base B

Nous allons étudier la traduction dans l'écriture matricielle des opérations que nous avons définies dans un espace vectoriel.

Somme de deux vecteurs

V est représenté par : $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$ W par : $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}$

Cela signifie que $V = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n$

et que $W = \beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_n V_n$

La somme de V et W est le vecteur S que nous savons parfaitement construire :

$$S = V + W = (\alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n) + (\beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_n V_n)$$

En nous servant de l'associativité, cela devient :

$$S = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n + \beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_n V_n$$

Par commutativité :

$$S = \alpha_1 V_1 + \beta_1 V_1 + \alpha_2 V_2 + \beta_2 V_2 + \ldots + \alpha_n V_n + \beta_n V_n$$

Par associativité :

$$S = (\alpha_1 V_1 + \beta_1 V_1) + (\alpha_2 V_2 + \beta_2 V_2) + \ldots + (\alpha_n V_n + \beta_n V_n)$$

Par distributivité de la multiplication par un nombre par rapport à l'addition des nombres :

$$S = (\alpha_1 + \beta_1) V_1 + (\alpha_2 + \beta_2) V_2 + \ldots + (\alpha_n + \beta_n) V_n$$

S est représenté par le système $(\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$ dans la base $(V_1, V_2, \ldots, V_n)$.
En écriture matricielle, cela se traduit par :

$$V = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} \qquad W = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix} \qquad S = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \ldots \\ \alpha_n + \beta_n \end{pmatrix}$$

et nous dirons que $\begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \ldots \\ \alpha_n + \beta_n \end{pmatrix}$ est la somme matricielle de V et W.

Soit :

$$\boxed{\begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \ldots \\ \alpha_n + \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix}}$$

Si nous transposons les propriétés du calcul vectoriel au calcul matriciel, nous voyons apparaître les propriétés suivantes :
$V + W = W + V$ (commutativité de l'addition vectorielle) devient :

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$$

(commutativité de l'addition matricielle)

$V + (W + Z) = (V + W) + Z$: l'associativité de l'addition vectorielle donne l'associativité de l'addition matricielle.

Le vecteur élément neutre O a pour matrice $\begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}$; cette matrice sera la matrice élément neutre pour l'addition matricielle :

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}$$

Il existe un inverse pour tout vecteur. Si

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \dots + \alpha_n V_n,$$

son opposé est $-V = (-\alpha_1) V_1 + (-\alpha_2) V_2 + \dots + (-\alpha_n) V_n$.

Écriture matricielle : $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ \dots \\ -\alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}$

Nous pouvons de même étudier la traduction des propriétés de la multiplication par un nombre réel dans le calcul matriciel.

Si $V = \alpha_1 V_1 + \alpha_2 V_2 + \dots + \alpha_n V_n$

alors $\alpha . V = \alpha . (\alpha_1 V_1 + \alpha_2 V_2 + \dots + \alpha_n V_n)$

$= (\alpha\alpha_1) V_1 + (\alpha\alpha_2) V_2 + \dots + (\alpha\alpha_n) V_n$

Si V est représenté par $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}$, αV est représenté par $\begin{pmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \\ \dots \\ \alpha\alpha_n \end{pmatrix}$.

Nous poserons $\alpha . \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \\ \dots \\ \alpha\alpha_n \end{pmatrix}$ et nous avons $1 . \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}$

La distributivité de la multiplication des nombres par rapport à l'addition des vecteurs : $\alpha(V_1 + V_2) = \alpha V_1 + \alpha V_2$ se traduit par :

$$\alpha . \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \dots \\ \alpha_n + \beta_n \end{pmatrix} = \alpha . \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} + \alpha . \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{pmatrix}$$

soit distributivité de la multiplication des nombres par rapport à l'addition des matrices.

La distributivité de la multiplication des nombres par rapport à l'addition des nombres : $(\alpha + \beta)\,V = \alpha V + \beta V$ donne une propriété analogue pour les matrices.
Enfin, l'associativité de la multiplication par un nombre s'écrit :

$$(\alpha \times \beta)\,V = \alpha.(\beta V)$$

$$(\alpha \times \beta).\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_n \end{pmatrix} = \alpha.\begin{pmatrix} \beta\alpha_1 \\ \beta\alpha_2 \\ \cdots \\ \beta\alpha_n \end{pmatrix} = \begin{pmatrix} \alpha\beta\alpha_1 \\ \alpha\beta\alpha_2 \\ \cdots \\ \alpha\beta\alpha_n \end{pmatrix}$$

Toutes ces propriétés donne à l'ensemble des matrices de nombres réels à n lignes et une colonne une structure d'espace vectoriel.

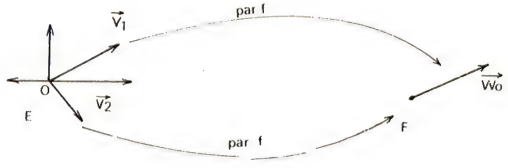
APPLICATIONS LINEAIRES

Nous allons maintenant nous intéresser aux applications qui nous permettent de passer d'un espace vectoriel E dans un espace vectoriel F.

APPLICATION LINÉAIRE OU HOMOMORPHISME D'ESPACE VECTORIEL

Nous notons E l'espace de départ et F l'espace d'arrivée. Nous supposons qu'ils sont tous les deux munis d'une structure d'espace vectoriel sur un même corps K. Une application de E dans F est la donnée d'un processus qui à tout vecteur $\vec{V}$ de l'espace de départ E nous permet de faire correspondre un vecteur $\vec{W}$ de l'espace d'arrivée F. Puisque E et F sont munis d'une structure particulière (structure d'espace vectoriel), il est naturel de se poser la question : l'application f de E dans F a-t-elle des propriétés particulières relatives à ces structures ?
Si f est une application tout à fait quelconque, nous ne pouvons pas mettre en évidence des propriétés particulières. Ainsi, supposons que E soit l'ensemble des vecteurs $\vec{V}$ d'origine O dans le plan, ainsi que F. Appelons f l'application qui à tout vecteur $\vec{V}$ de E associe toujours le même vecteur $\vec{W_0}$ de F. Cette application n'a pas de propriétés intéressantes.

$$\forall \vec{V} \in E, \quad f(\vec{V}) = \vec{W_0}$$



Par contre, une propriété intéressante serait la suivante : E et F sont munis d'une opération interne appelée addition.
Soient V_1 et V_2 deux vecteurs de E, $f(V_1)$ et $f(V_2)$ leur image par f dans F. Il serait intéressant, lorsqu'on cherche l'image du vecteur $V_1 + V_2$ par f, que ce soit la somme des images $f(V_1)$ et $f(V_2)$, c'est-à-dire que l'image de la somme de deux vecteurs V_1 et V_2 soit la somme des images $f(V_1)$ et $f(V_2)$.
De même, pour la deuxième loi (multiplication par un scalaire), il serait intéressant que l'image du vecteur αV soit égale à α fois l'image $f(V)$.
Dans ce cas, on dit que f respecte les structures et on l'appelle *homomorphisme d'espace vectoriel* (application d'un espace vectoriel dans un espace vectoriel respectant leur structure) ou *application linéaire*.

Une application f de E dans F espace vectoriel sur le même corps K est appelée application linéaire si et seulement si :

$$\forall V_1 \text{ et } V_2 \in E \qquad f(V_1 + V_2) = f(V_1) + f(V_2)$$
$$\forall V \in E, \quad \forall \alpha \in K \qquad f(\alpha V) = \alpha f(V)$$

L'homothétie est une application linéaire

Supposons que E et F soient l'espace vectoriel des vecteurs du plan d'origine O. Définissons ainsi l'application f : à tout vecteur V de E on associe le vecteur kV, où k est un nombre réel. Nous avons bien kV dans F.
Regardons si les deux propriétés précédentes sont vérifiées :

$$V_1 \longmapsto kV_1 = W_1$$
$$V_2 \longmapsto kV_2 = W_2$$
$$(V_1 + V_2) \longmapsto k(V_1 + V_2) = W_3$$

Et, grâce à la distributivité de la multiplication par un nombre réel par rapport à la somme des vecteurs, nous avons :

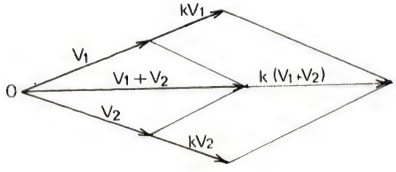
$$W_3 = k(V_1 + V_2) = kV_1 + kV_2 = W_1 + W_2$$

De même, grâce à l'associativité de la multiplication par un scalaire, nous avons :

$$V \longmapsto kV$$

$$\alpha V \longmapsto k(\alpha V) \quad \text{et} \quad k(\alpha V) = (k\alpha)\,V = (\alpha k)\,V = \alpha(kV)$$

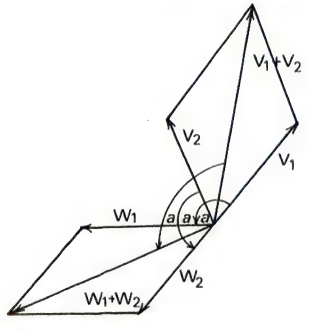
Cette application est linéaire. Or elle est bien connue. C'est une homothétie de centre O et de rapport k.



L'homothétie de centre O *est une application linéaire de l'ensemble des vecteurs d'origine* O *sur eux-mêmes.*

La rotation est une application linéaire

Nous pouvons également considérer l'application suivante : à tout vecteur V de E, nous associons le vecteur W faisant un angle a avec V et de même longueur.
A V_1, nous associons W_1 qui fait un angle a avec V_1 et de même longueur que V_1.
A V_2, nous associons W_2 qui fait un angle a avec V_2 et de même longueur que V_2.
A $V_1 + V_2$, nous associons W_3, qui fait un angle a avec $V_1 + V_2$ et de même longueur que $V_1 + V_2$.
Or ce vecteur W_3 est bien la somme de W_1 et W_2.
De même, $f(\alpha V_1)$ est un vecteur de même longueur que αV_1 et faisant un angle a avec αV_1. Il est identique au vecteur αW_1, où W_1 est un vecteur faisant un angle a avec V_1 et de même longueur.



Cette application est la rotation d'angle a. La rotation d'angle a est une application linéaire de l'ensemble des vecteurs d'origine O sur eux-mêmes.

Il ne faudrait pas croire que toute application d'un espace vectoriel E *dans un espace vectoriel* F *soit linéaire.*

Ainsi, nous avons vu que l'ensemble $\mathbb{R} \times \mathbb{R}$ des couples de nombres réels pouvait être muni d'une structure d'espace vectoriel, suivant les lois suivantes :

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2,\ b_1 + b_2)$$
$$\alpha(a_1, b_1) = (\alpha a_1, \alpha b_1)$$

Considérons l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$, qui à tout couple (a, b) associe le couple $\left(\frac{1}{a}, \frac{1}{b}\right)$, si a et b sont différents de 0 et à $(0, b)$ associe $\left(0, \frac{1}{b}\right)$, à $(a, 0)$ associe $\left(\frac{1}{a}, 0\right)$.
Cette application est parfaitement définie de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Vérifie-t-elle $f(V_1 + V_2) = f(V_1) + f(V_2)$?

Étudions d'abord les vecteurs vérifiant $a_1 \neq 0$, $b_1 \neq 0$, $a_2 \neq 0$, $b_2 \neq 0$ et $(a_1 + a_2) \neq 0$, $(b_1 + b_2) \neq 0$.

$$\vec{V_1} = (a_1, b_1) \qquad f(V_1) = \left(\frac{1}{a_1}, \frac{1}{b_1}\right)$$
$$\vec{V_2} = (a_2, b_2) \qquad f(V_2) = \left(\frac{1}{a_2}, \frac{1}{b_2}\right)$$
$$V_1 + V_2 = (a_1 + a_2,\ b_1 + b_2)$$
$$f(V_1 + V_2) = \left(\frac{1}{a_1 + a_2}, \frac{1}{b_1 + b_2}\right).$$

Nous n'avons pas :

$$f(V_1 + V_2) = f(V_1) + f(V_2)$$

car nous n'avons pas :

$$\left(\frac{1}{a_1 + a_2}, \frac{1}{b_1 + b_2}\right) = \left[\left(\frac{1}{a_1}, \frac{1}{b_1}\right) + \left(\frac{1}{a_2}, \frac{1}{b_2}\right)\right]$$
$$= \left(\frac{1}{a_1} + \frac{1}{a_2}, \frac{1}{b_1} + \frac{1}{b_2}\right).$$

Ce n'est pas la peine de considérer d'autres vecteurs puisque pour ceux-ci, déjà, cette application n'est pas linéaire.
L'application ne vérifie pas les propriétés de définition des applications linéaires.

Forme linéaire

Nous avons vu que l'ensemble des nombres réels $\mathbb{R}$ pouvait être considéré comme un espace vectoriel sur lui-même. Par conséquent, il est tout à fait possible de définir une application linéaire f d'un

espace vectoriel E dans $\mathbb{R}$ (il est naturellement nécessaire que le corps des scalaires pour E soit l'ensemble des nombres réels). Une telle application porte un nom particulier : on appelle forme linéaire une application linéaire d'un espace vectoriel E (dont le corps des scalaires est $\mathbb{R}$) sur l'ensemble des nombres réels.

Propriétés des applications linéaires

1) *L'image par une application linéaire du vecteur nul de* E *est le vecteur nul de* F.
Soient E et F deux espaces vectoriels définis sur un même corps K et f une application linéaire de E dans F :

$$f(O_E) = O_F.$$

En effet, $f(x + O) = f(x) + f(O)$ d'après 1).
Or $x + O = x \Longrightarrow f(x) = f(x) + f(O)$.
Par conséquent, $f(O) = O_F$.

2) *L'image d'un système générateur de* E *est un système générateur de* f(E).

Soit $(V_1, V_2, \ldots, V_n)$ un système générateur de E. Tout vecteur de E peut s'écrire comme combinaison linéaire des n vecteurs $(V_1, V_2, \ldots, V_n)$. :

$$\forall\, V \in E, \exists\, \alpha_1, \alpha_2, \ldots, \alpha_n \text{ tel que}$$
$$V = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n.$$

Dans ce cas, $f(V) = \alpha_1 f(V_1) + \alpha_2 f(V_2) + \ldots + \alpha_n f(V_n)$, à cause de la linéarité de l'application f. Tout vecteur de l'espace image f(E) peut donc s'exprimer comme combinaison linéaire des n vecteurs $f(V_1), f(V_2), \ldots, f(V_n)$. Nous pouvons donc bien dire que $f(V_1)$, $f(V_2), \ldots, f(V_n)$ engendre f(E).

3) *L'image réciproque d'un système libre est un système libre.*
Soit $W_1, W_2, \ldots, W_p$ un système libre de f(E). Cela signifie que l'égalité $\alpha_1 W_1 + \alpha_2 W_2 + \ldots + \alpha_p W_p = O$ ne peut être réalisée que si $\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0$. On peut démontrer dans ce cas que les vecteurs $V_1, V_2, \ldots, V_p$ vérifiant $f(V_1) = W_1, f(V_2) = W_2, \ldots, f(V_p) = W_p$ forment un système libre. C'est-à-dire qu'on démontre que, pour arriver à l'égalité $\beta_1 V_1 + \beta_2 V_2 + \ldots + \beta_p V_p = O$, il faut nécessairement que $\beta_1 = \beta_2 = \ldots = \beta_p = 0$. Et, puisque l'image réciproque d'un système libre est un système libre, nous pouvons affirmer que l'image d'un système lié est un système lié. En effet, si l'image d'un système lié était un système libre, dans ce cas nous aurions un système libre dont l'image réciproque est un système lié, ce qui est en contradiction avec la propriété que nous venons de démontrer.

Cela nous permet d'affirmer qu'une application linéaire n'augmente jamais la dimension des espaces vectoriels. Par contre, elle peut la diminuer.

4) *L'image réciproque du vecteur nul de* F *est un sous-espace vectoriel de* E. On l'appelle *noyau de l'application f.* On le note Ker (f).
Cela signifie que c'est un sous-ensemble de E ayant lui-même une structure d'espace vectoriel ou (ce qui est équivalent ◆) qui est stable pour la loi d'addition et la loi de multiplication par un scalaire.

◆ Nous n'en donnons pas la démonstration ici. Elle est néanmoins assez facile. Pour cela, on procède de la même façon que pour le sous-groupe (voir l'article « l'Algèbre », page 35).

Exemples de recherche du noyau d'une application linéaire.
Soit $\mathfrak{D}$ l'ensemble des fonctions définies sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$, qui sont dérivables sur $\mathbb{R}$ tout entier. Grâce aux propriétés de la dérivation, le lecteur pourrait montrer que cet ensemble $\mathfrak{D}$ a une structure d'espace vectoriel sur $\mathbb{R}$.
Considérons l'application d définie de $\mathfrak{D}$ dans $\mathbb{R}$ qui, à toute application de $\mathfrak{D}$, associe sa dérivée au point c :

$$d : f \in \mathfrak{D} \longmapsto f'(c) \in \mathbb{R}.$$

Le noyau est formé de l'ensemble (F_c) des fonctions f de $\mathfrak{D}$ tel que $f'(c) = 0$. On peut démontrer directement que l'ensemble (F_c) a une structure d'espace vectoriel.
Par contre, si l'application linéaire considérée est la suivante :

$$f \in \mathfrak{D} \longmapsto f' \in \mathfrak{F},$$

à toute fonction f de $\mathfrak{D}$ elle associe sa fonction dérivée qui est une fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

L'élément neutre de $\mathfrak{F}$ est l'application $f(x) = 0$, $\forall\, x \in \mathbb{R}$, l'application identiquement nulle. Le noyau est donc formé de l'ensemble des fonctions $f(x) \in \mathfrak{D}$ telles que $f'(x) = 0$, $\forall\, x \in \mathbb{R}$. C'est l'ensemble des fonctions à dérivée identiquement nulle. C'est l'ensemble $\mathfrak{G}$ des fonctions constantes de $\mathfrak{D}$.

Isomorphismes d'espaces vectoriels
Les applications linéaires bijectives ◆ d'un espace vectoriel E dans un espace vectoriel F forment un ensemble particulier d'applications linéaires.
Par définition, ce sont les applications linéaires qui sont à la fois surjectives $\big(f(E) = F\big)$ et injectives $\big(\text{si } V_1 \neq V_2, \text{ alors } f(V_1) \neq f(V_2)\big)$.
On les appelle *isomorphismes d'espace vectoriel.*

◆ Voir l'article « les Ensembles », page 278.

Elles jouissent de propriétés particulières :
— le noyau Ker f est réduit au vecteur nul de E ;
— l'application inverse définie sur F à valeurs dans E est un isomorphisme ;
— si E et F sont de dimensions finies, alors :

$$\dim E = \dim F$$

ÉCRITURE MATRICIELLE DES APPLICATIONS LINÉAIRES

Nous avons vu que, dans le cas particulier d'un espace vectoriel de dimension finie n, nous pouvons exprimer tout vecteur de cet espace à partir de n vecteurs $(V_1, V_2, V_3, \ldots, V_n)$ qui en formaient une base. Regardons, dans le cas particulier où les deux espaces E et F sont de dimension finie, comment on peut traduire une application linéaire f.

Supposons donc que E soit de dimension finie n avec la base $B = (V_1, V_2, \ldots, V_n)$ et que F soit de dimension finie p avec la base $B' = (W_1, W_2, \ldots, W_p)$, E et F étant toujours définis sur le même corps de référence K. Supposons aussi que nous connaissions l'image par f de la base B. Que pouvons-nous dire alors de l'image par f d'un vecteur V quelconque de E?

Nous savons que nous pouvons exprimer V comme combinaison linéaire des n vecteurs $(V_1, V_2, \ldots, V_n)$, c'est-à-dire que :

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n.$$

D'après les propriétés des applications linéaires, nous avons alors :

$$f(V) = \alpha_1 f(V_1) + \alpha_2 f(V_2) + \ldots + \alpha_n f(V_n).$$

Nous voyons que si nous connaissons $f(V_1)$, $f(V_2)$, $\ldots$, $f(V_n)$, alors $f(V)$ est parfaitement déterminée.

> Une application linéaire d'un espace vectoriel E de base $B = (V_1, V_2, \ldots, V_n)$ dans un espace vectoriel F est parfaitement déterminée si nous connaissons l'image par f de chaque vecteur de la base B.

Mais comment pouvons-nous exprimer les images des vecteurs $f(V_1)$, $f(V_2)$, $\ldots$, $f(V_n)$ dans F? Si F est de dimension p avec la base $B' = (W_1, W_2, \ldots, W_p)$, nous pouvons décomposer chaque vecteur $f(V_1), f(V_2), \ldots, f(V_n)$ sur la base B'.

$f(V_1)$ s'exprimera ainsi :

$$f(V_1) = a_1 W_1 + a_2 W_2 + \ldots + a_p W_p$$

De même :

$$f(V_2) = b_1 W_1 + b_2 W_2 + \ldots + b_p W_p$$

et ainsi de suite.

En général, on adopte la convention suivante :

$$\begin{aligned}
f(V_1) &= \alpha_{1,1} W_1 + \alpha_{2,1} W_2 + \ldots + \alpha_{p,1} W_p \\
f(V_2) &= \alpha_{1,2} W_1 + \alpha_{2,2} W_2 + \ldots + \alpha_{p,2} W_p \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
f(V_n) &= \alpha_{1,n} W_1 + \alpha_{2,n} W_2 + \ldots + \alpha_{p,n} W_p
\end{aligned}$$

où α_{ij} désigne la composante du vecteur $f(V_j)$ sur le vecteur W_i.

Par commodité, on inscrit ces coefficients dans un tableau rectangulaire où, par convention, on note en colonne i les coefficients du vecteur $f(V_i)$ sur la base $W_1, W_2, \ldots, W_p$.

	$f(V_1)$	$f(V_2)$		$f(V_n)$	
	$\alpha_{1,1}$	$\alpha_{1,2}$		$\alpha_{1,n}$	$/W_1$
	$\alpha_{2,1}$	$\alpha_{2,2}$		$\alpha_{2,n}$	$/W_2$
	$\alpha_{3,1}$	$\alpha_{3,2}$		$\alpha_{3,n}$	$/W_3$
M =	.	.		.	.
	.	.		.	.
	.	.		.	.
	$\alpha_{p,1}$	$\alpha_{p,2}$		$\alpha_{p,n}$	$/W_p$

Ce tableau rectangulaire est appelé matrice M de l'application linéaire f. C'est un tableau à p lignes et n colonnes ; α_{ij} est un élément de la matrice M.

Pouvons-nous utiliser cette notation de façon pratique ? Nous avons déjà introduit l'écriture matricielle pour noter les coefficients d'un vecteur V sur une base B. Nous l'inscrivions en une matrice unicolonne. Nous allons maintenant chercher à exprimer $f(V)$ dans une écriture matricielle et nous verrons si cela permet de munir l'ensemble des matrices de nouvelles propriétés.

Nous savons que, dans la base B, V peut être représenté par la matrice de ses composantes : $(V) = \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}$ dans la base B.

De plus, nous avons :

$$f(V) = a_1 f(V_1) + a_2 f(V_2) + \ldots + a_n f(V_n)$$

Supposons que

$$f(V_1) = \alpha_{1,1} W_1 + \alpha_{2,1} W_2 + \ldots + \alpha_{p,1} W_p$$
$$f(V_2) = \alpha_{1,2} W_1 + \alpha_{2,2} W_2 + \ldots + \alpha_{p,2} W_p$$
etc.

Alors, nous avons :

$$\begin{aligned} f(V) = {} & a_1[\alpha_{1,1} W_1 + \alpha_{2,1} W_2 + \ldots + \alpha_{p,1} W_p] \\ & + a_2[\alpha_{1,2} W_1 + \alpha_{2,2} W_2 + \ldots + \alpha_{p,2} W_p] \\ & + \ldots \\ & + a_n[\alpha_{1,n} W_1 + \alpha_{2,n} W_2 + \ldots + \alpha_{p,n} W_p] \end{aligned}$$

soit

$$\begin{aligned} & = (a_1\alpha_{1,1} + a_2\alpha_{1,2} + \ldots + a_n\alpha_{1,n})\, W_1 \\ & + (a_1\alpha_{2,1} + a_2\alpha_{2,2} + \ldots + a_n\alpha_{2,n})\, W_2 \\ & + \ldots \\ & + (a_1\alpha_{p,1} + a_2\alpha_{p,2} + \ldots + a_n\alpha_{p,n})\, W_p \end{aligned}$$

Nous voyons que, dans la base B′ de F, le vecteur $f(V)$ a les coefficients notés ci-dessus.
Nous pouvons donc le représenter par la matrice unicolonne :

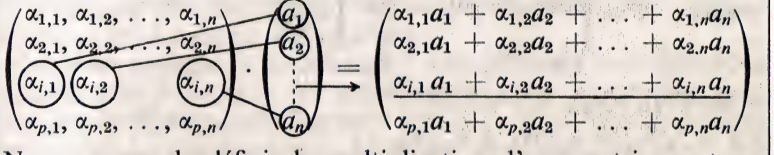
$$(Y) = (f(V)) = \begin{pmatrix} a_1\alpha_{1,1} + a_2\alpha_{1,2} + \ldots + a_n\alpha_{1,n} \\ a_1\alpha_{2,1} + a_2\alpha_{2,2} + \ldots + a_n\alpha_{2,n} \\ \ldots \\ a_1\alpha_{p,1} + a_2\alpha_{p,2} + \ldots + a_n\alpha_{p,n} \end{pmatrix}$$

dans la base B′.

Poussons l'écriture conventionnelle jusqu'au bout.
Nous avons l'écriture fonctionnelle $f.(V) = Y$.
Nous la traduisons par l'écriture matricielle $M.(V) = (Y)$.
C'est-à-dire que, par convention, nous poserons :

$$\begin{pmatrix} \alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{1,n} \\ \alpha_{2,1}, \alpha_{2,2}, \ldots, \alpha_{2,n} \\ \ldots \\ \alpha_{p,1}, \alpha_{p,2}, \ldots, \alpha_{p,n} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}a_1 + \alpha_{1,2}a_2 + \ldots + \alpha_{1,n}a_n \\ \alpha_{2,1}a_1 + \alpha_{2,2}a_2 + \ldots + \alpha_{2,n}a_n \\ \ldots \\ \alpha_{p,1}a_1 + \alpha_{p,2}a_2 + \ldots + \alpha_{p,n}a_n \end{pmatrix}$$

Le produit d'une matrice M rectangulaire à p lignes et n colonnes par une matrice (V) à n lignes et une colonne est une matrice (Y) à p lignes et une colonne. On obtient le terme de la ligne j de Y en multipliant terme à terme les éléments de la ligne j de M par ceux de la matrice V, et on ajoute ces produits. C'est-à-dire qu'on a le schéma :



Nous venons de définir la multiplication d'une matrice rectangulaire à p lignes et n colonnes par une matrice unicolonne à n lignes. Le produit en est une matrice unicolonne à p lignes.

Nous pouvons pousser plus loin la représentation des applications linéaires par des matrices. Comme nous l'avons fait quand nous avons introduit la représentation matricielle des vecteurs, nous allons transporter sur les matrices les opérations définies sur les applications linéaires. Nous noterons rapidement les opérations que nous pouvons définir sur l'ensemble des applications linéaires. Nous laisserons au lecteur le soin de faire les démonstrations des propriétés énoncées ; ces démonstrations sont en général extrêmement simples.

Addition

Soient f_1 et f_2 deux applications linéaires définies sur E à valeurs dans F. Nous appelons somme de f_1 et f_2 l'application s de E dans F ainsi définie : à tout vecteur V de E, elle associe le vecteur $f_1(V) + f_2(V)$ de F.
Cette application s de E dans F est une application linéaire.
Nous supposerons f_1 déterminée par l'image qu'elle prend sur les n vecteurs de base $V_1, V_2, \ldots, V_n$. C'est-à-dire que nous connaissons :

$$\begin{aligned}
f_1(V_1) &= \alpha_{1,1}W_1 + \alpha_{2,1}W_2 + \ldots + \alpha_{p,1}W_p \\
f_1(V_2) &= \alpha_{1,2}W_1 + \alpha_{2,2}W_2 + \ldots + \alpha_{p,2}W_p \\
&\ldots \\
f_1(V_n) &= \alpha_{1,n}W_1 + \alpha_{2,n}W_2 + \ldots + \alpha_{p,n}W_p
\end{aligned}$$

représentée par la matrice $M = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \ldots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \ldots & \alpha_{2,n} \\ \ldots & & & \\ \alpha_{p,1} & \alpha_{p,2} & \ldots & \alpha_{p,n} \end{pmatrix}$

De même, f_2 est connue par l'image des vecteurs de base :

$$\begin{aligned}
f_2(V_1) &= \beta_{1,1}W_1 + \beta_{2,1}W_2 + \ldots + \beta_{p,1}W_p \\
f_2(V_2) &= \beta_{1,2}W_1 + \beta_{2,2}W_2 + \ldots + \beta_{p,2}W_p \\
&\ldots \\
f_2(V_n) &= \beta_{1,n}W_1 + \beta_{2,n}W_2 + \ldots + \beta_{p,n}W_p
\end{aligned}$$

représentée par la matrice $N = \begin{pmatrix} \beta_{1\,1}, & \beta_{1,2}, & \ldots, & \beta_{1,n} \\ \beta_{2,1}, & \beta_{2,2}, & \ldots, & \beta_{2,n} \\ \ldots & & & \\ \beta_{p,1}, & \beta_{p,2}, & \ldots, & \beta_{p,n} \end{pmatrix}$

Quelle est la valeur de $s(V_1), s(V_2), \ldots, s(V_n)$?

$$\begin{aligned}
s(V_1) = f_1(V_1) + f_2(V_1) &= \alpha_{1,1}W_1 + \alpha_{2,1}W_2 \\
&+ \ldots + \alpha_{p,1}W_p + \beta_{1,1}W_1 + \beta_{2,1}W_2 + \ldots + \beta_{p,1}W_p \\
&= (\alpha_{1,1} + \beta_{1,1})\, W_1 + (\alpha_{2,1} + \beta_{2,1})\, W_2 \\
&+ \ldots + (\alpha_{p,1} + \beta_{p,1})\, W_p
\end{aligned}$$

Et nous pouvons faire le même calcul pour les autres vecteurs.

Nous représentons donc s par la matrice S suivante :

$$S = \begin{pmatrix} \alpha_{1,1} + \beta_{1,1}, & \alpha_{1,2} + \beta_{1,2}, & \alpha_{1,3} + \beta_{1,3}, & \ldots, & \alpha_{1,n} + \beta_{1,n} \\ \ldots & & & & \\ \alpha_{p,1} + \beta_{p,1}, & \alpha_{p,2} + \beta_{p,2}, & \ldots, & & \alpha_{p,n} + \beta_{p,n} \end{pmatrix}$$

Et l'écriture fonctionnelle $f_1 + f_2 = s$ se traduit par l'écriture matricielle $M + N = S$, soit

$$\begin{pmatrix} \alpha_{1,1}, \alpha_{1,2}, \alpha_{1,3}, \ldots, \alpha_{1,n} \\ \ldots \\ \alpha_{p,1}, \alpha_{p,2}, \alpha_{p,3}, \ldots, \alpha_{p,n} \end{pmatrix} + \begin{pmatrix} \beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,n} \\ \ldots \\ \beta_{p,1}, \beta_{p,2}, \ldots, \beta_{p,n} \end{pmatrix} = \begin{pmatrix} \alpha_{1,1} + \beta_{1,1}, \ldots, \alpha_{1,n} + \beta_{1,n} \\ \ldots \\ \alpha_{p,1} + \beta_{p,1}, \ldots, \alpha_{p,n} + \beta_{p,n} \end{pmatrix}$$

La somme de deux matrices rectangulaires à p lignes et n colonnes est une matrice à p lignes et n colonnes. L'élément de ligne i et de colonne j dans cette matrice somme est la somme des éléments de ligne i et de colonne j des deux autres matrices.

On pourrait montrer que l'addition de deux applications linéaires d'un espace vectoriel E dans un espace vectoriel F est commutative et associative. Il en sera de même de l'addition des matrices rectangulaires à p lignes et n colonnes.
Il existe une application linéaire de E dans F élément neutre pour l'addition des applications linéaires. C'est l'application qui à tout vecteur V de E associe le vecteur nul de F. Nous la noterons f_0.

$$f_0 : E \longrightarrow F$$
$$V \longmapsto O_F \text{ pour tout V.}$$

Elle est représentée par la matrice suivante appelée matrice nulle à p lignes et n colonnes :

$$(O) = \begin{pmatrix} 0, \ldots, 0 \\ \ldots \\ 0, \ldots, 0 \end{pmatrix} \text{ à } p \text{ lignes et } n \text{ colonnes.}$$

Cette matrice (O) sera élément neutre pour l'addition des matrices rectangulaires à p lignes et n colonnes.
Enfin, toute application linéaire de E dans F admet une application opposée pour l'addition des applications. C'est la fonction $-f$. Tous ses coefficients sont opposés à ceux de f. Si M représente la fonction f, alors la matrice $-$ M représentera l'application $-f$ et nous aurons (M) + (— M) = (O).

Toutes ces propriétés montrent que l'ensemble des matrices rectangulaires ayant le même nombre de lignes p et le même nombre de colonnes n est muni d'une structure de groupe commutatif pour l'addition.

Multiplication par un scalaire ◆

◆ Voir la définition du scalaire, ci-dessus, page 61.

Si nous prenons un scalaire α appartenant au corps de référence K commun aux deux espaces vectoriels E et F, alors nous savons

définir la multiplication de l'application linéaire f par le scalaire α ; c'est l'application qui à tout vecteur V de E associe le vecteur $\alpha \times f(V)$ dans F. Nous pouvons montrer que l'application αf est elle aussi une application linéaire de E dans F.
Voyons quels sont les coefficients de l'application αf dans la base B de E et B′ de F :

$$\begin{aligned} \alpha f(V_1) &= \alpha[\alpha_{1,1} W_1 + \alpha_{2,1} W_2 + \alpha_{3,1} W_3 + \ldots + \alpha_{p,1} W_p] \\ &= \alpha\alpha_{1,1} . W_1 + \alpha\alpha_{2,1} . W_2 + \ldots + \alpha\alpha_{p,1} . W_p \end{aligned}$$

et de même pour les autres vecteurs de la base B : V_2, V_3, ..., V_n.
Par conséquent, (αf) sera représentée par la matrice :

$$P = \begin{pmatrix} \alpha . \alpha_{1,1}, \alpha . \alpha_{1,2}, \ldots, \alpha . \alpha_{1,n} \\ \alpha . \alpha_{2,1}, \alpha . \alpha_{2,2}, \ldots, \alpha . \alpha_{2,n} \\ \ldots \\ \alpha . \alpha_{p,1}, \alpha . \alpha_{p,2}, \ldots, \alpha . \alpha_{p,n} \end{pmatrix}$$

Nous avons l'égalité fonctionnelle :

$$\alpha \times f = (\alpha f)$$

Elle se traduit par l'égalité matricielle :

$$\alpha . M = P$$

soit :

$$\alpha \begin{pmatrix} \alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{1,n} \\ \alpha_{2,1}, \alpha_{2,2}, \ldots, \alpha_{2,n} \\ \ldots \\ \alpha_{p,1}, \alpha_{p,2}, \ldots, \alpha_{p,n} \end{pmatrix} = \begin{pmatrix} \alpha\alpha_{1,1}, \alpha\alpha_{1,2}, \ldots, \alpha\alpha_{1,n} \\ \alpha\alpha_{2,1}, \alpha\alpha_{2,2}, \ldots, \alpha\alpha_{2,n} \\ \ldots \\ \alpha\alpha_{p,1}, \alpha\alpha_{p,2}, \ldots, \alpha\alpha_{p,n} \end{pmatrix}$$

Les différentes propriétés de la multiplication par un scalaire des applications linéaires de E dans F sont les suivantes : distributivité par rapport à l'addition des applications, par rapport à l'addition des scalaires de K ; associativité de la multiplication des scalaires ; l'élément neutre de la multiplication des scalaires est aussi l'élément neutre de cette opération externe. Nous pouvons les transposer pour la multiplication par un scalaire des matrices rectangulaires à p lignes et n colonnes ; distributivité par rapport à l'addition des matrices, par rapport à l'addition des scalaires de K ; associativité de la multiplication des scalaires ; un élément neutre de la multiplication des scalaires est aussi un élément neutre de cette opération externe dans l'ensemble des matrices rectangulaires à p lignes et n colonnes.
Ces deux lois, addition et multiplication par un élément de K, munissent l'ensemble des matrices à p lignes et n colonnes d'une structure d'espace vectoriel. Cela est lié au fait que l'ensemble des

applications linéaires définies sur un espace vectoriel E de dimension n à valeurs dans un espace vectoriel F de dimension p a lui aussi une structure d'espace vectoriel sur K (si K est le corps de référence commun aux deux espaces E et F).

Multiplication de deux matrices

Il existe encore une autre opération sur les applications que nous n'avons pas utilisée : c'est la loi de composition des applications. E, F et G désignent trois espaces vectoriels définis sur le même corps de référence K. Nous supposons que :

E est de dimension n et de base $B = (V_1, V_2, \ldots, V_n)$,

F de dimension p et de base $B' = (X_1, X_2, \ldots, X_p)$,

G de dimension q et de base $B'' = (Y_1, Y_2, \ldots, Y_q)$.

Alors l'application h définie dans E à valeurs dans G qui à tout vecteur V de E associe le vecteur Y de G ainsi défini :

$$Y = h(V) = g[f(V)] = g(Y) \text{ si } Y = f(V) \text{ vecteur de F,}$$

s'appelle application composée de f par g.

On peut montrer facilement que l'application h ainsi définie est une application linéaire de E dans G. Nous avons l'égalité fonctionnelle :

$$h = g \circ f$$

Voyons comment elle est traduite par les matrices.

Supposons que

$$f(V_1) = \alpha_{1,1} X_1 + \alpha_{2,1} X_2 + \ldots + \alpha_{p,1} X_p$$

etc.

D'où

$$M = \begin{pmatrix} \alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{1,n} \\ \alpha_{2,1}, \alpha_{2,2}, \ldots, \alpha_{2,n} \\ \ldots \\ \alpha_{p,1}, \alpha_{p,2}, \ldots, \alpha_{p,n} \end{pmatrix}$$

Supposons que

$$g(X_1) = \beta_{1,1} Y_1 + \beta_{2,1} Y_2 + \ldots + \beta_{q,1} Y_p$$

etc.

D'où

$$N = \begin{pmatrix} \beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,p} \\ \beta_{2,1}, \beta_{2,2}, \ldots, \beta_{2,p} \\ \ldots \\ \beta_{q,1}, \beta_{q,2}, \ldots, \beta_{q,p} \end{pmatrix}$$

Évaluons les coefficients de l'application $h = g \circ f$:

$$\begin{aligned} h(V_1) = g[f(V_1)] &= g[\alpha_{1,1} X_1 + \alpha_{2,1} X_2 + \ldots + \alpha_{p,1} X_p] \\ &= \alpha_{1,1}\, g(X_1) + \alpha_{2,1}\, g(X_2) + \ldots + \alpha_{p,1}\, g(X_p) \\ &= \alpha_{1,1}\,[\beta_{1,1} Y_1 + \beta_{2,1} Y_2 + \ldots + \beta_{q,1} Y_q] \\ &\quad + \alpha_{2,1}\,[\beta_{1,2} Y_1 + \beta_{2,2} Y_2 + \ldots + \beta_{q,2} Y_q] \\ &\quad + \ldots \end{aligned}$$

$$\begin{aligned} h(V_1) = {} & (\alpha_{1,1}\beta_{1,1} + \alpha_{2,1}\beta_{1,2} + \ldots + \alpha_{p,1}\beta_{1,p})\, Y_1 \\ & + (\alpha_{1,1}\beta_{2,1} + \alpha_{2,1}\beta_{2,2} + \ldots + \alpha_{p,1}\beta_{2,p})\, Y_2 \\ & + \ldots \\ & + (\alpha_{1,1}\beta_{q,1} + \alpha_{2,1}\beta_{q,2} + \ldots + \alpha_{p,1}\beta_{q,p})\, Y_q \end{aligned}$$

Nous pouvons, de la même façon, évaluer $h(V_2)$, $h(V_3)$, ..., $h(V_n)$. Soit H la matrice de l'application h. Nous avons l'égalité matricielle :

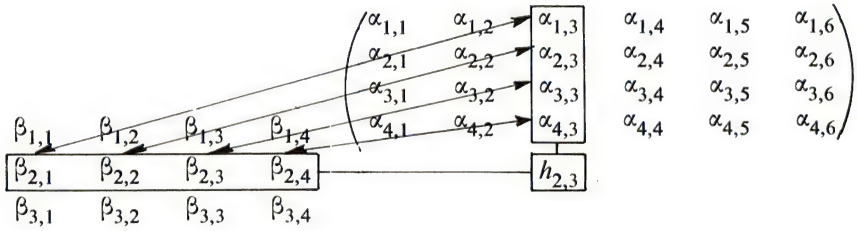
$$N \times M = H$$

Les coefficients se calculent de la façon suivante :

$$\begin{pmatrix} \beta_{1,1}, \beta_{1,2}, \ldots, \beta_{1,p} \\ \beta_{2,1}, \beta_{2,2}, \ldots, \beta_{2,p} \\ \ldots \\ \beta_{q,1}, \beta_{q,2}, \ldots, \beta_{q,p} \end{pmatrix} \begin{pmatrix} \alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{1,n} \\ \alpha_{2,1}, \alpha_{2,2}, \ldots, \alpha_{2,n} \\ \ldots \\ \alpha_{p,1}, \alpha_{p,2}, \ldots, \alpha_{p,n} \end{pmatrix}$$

$$= \begin{pmatrix} \beta_{1,1}\alpha_{1,1} + \beta_{1,2}\alpha_{2,1} + \ldots + \beta_{1,p}\alpha_{p,1}, \ldots, \beta_{1,1}\alpha_{1,n} + \beta_{1,2}\alpha_{2,n} + \ldots + \beta_{1,p}\alpha_{p,n} \\ \beta_{2,1}\alpha_{1,1} + \beta_{2,2}\alpha_{2,1} + \ldots + \beta_{2,p}\alpha_{p,1}, \ldots, \beta_{2,1}\alpha_{1,n} + \beta_{2,2}\alpha_{2,n} + \ldots + \beta_{2,p}\alpha_{p,n} \\ \ldots \\ \beta_{q,1}\alpha_{1,1} + \beta_{q,2}\alpha_{2,1} + \ldots + \beta_{q,p}\alpha_{p,1}, \ldots, \beta_{q,1}\alpha_{1,n} + \beta_{q,2}\alpha_{2,n} + \ldots + \beta_{q,p}\alpha_{p,n} \end{pmatrix}.$$

Cela permet de définir la multiplication d'une matrice à q lignes et p colonnes par une matrice à p lignes et n colonnes. Le produit est une matrice à q lignes et n colonnes. On obtient le terme de la i-ième ligne et de la j-ième colonne en multipliant terme à terme les éléments de la i-ième ligne de la première matrice et ceux de j-ième colonne de la deuxième matrice, et on ajoute ces produits, c'est-à-dire qu'on a le schéma suivant :



On multiplie :
— le premier élément de la *i*-ième ligne de N par le premier élément de la *j*-ième colonne de M ;
— le deuxième élément de la *i*-ième ligne de N par le deuxième élément de la *j*-ième colonne de M ;
— le troisième élément de la *i*-ième ligne de N par le troisième élément de la *j*-ième colonne de B, et ainsi de suite;

Ce qui s'écrit symboliquement :

$$h_{ij} = \sum_{1 \leqslant k \leqslant q} \beta_{ik}\, \alpha_{kj}$$

La multiplication d'applications étant associative, il en sera de même de la multiplication des matrices. Mais attention : il y a des conditions de dimension indispensables à remplir.
En général, la multiplication de deux matrices n'est pas commutative : attention à l'ordre de l'écriture.

PETIT HISTORIQUE ET UTILISATION DES MATRICES

Nous venons de rencontrer les matrices comme représentant des vecteurs et des applications linéaires. L'écriture matricielle et les opérations que nous venons de définir peuvent servir dans des domaines beaucoup plus vastes. Elles peuvent même devenir dans ce cas des outils extrêmement pratiques et riches, comme nous allons le voir maintenant.

En fait, une matrice est un tableau rectangulaire ou carré, qui peut être rempli de nombres, de lettres ou de toute autre série de symboles. Chaque nombre se trouve à l'intersection d'une ligne et d'une colonne. On voit tout de suite que ce tableau permet une lecture suivant deux coordonnées donnant, l'une, la signification de chaque colonne (verticale), l'autre, la signification de chaque ligne (horizontale).

Différents exemples de matrices

×	0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9
2	0	2	4	6	8	10	12	14	16	18
3	0	3	6	9	12	15	18	21	24	27
4	0	4	8	12	16	20	24	28	32	36
5	0	5	10	15	20	25	30	35	40	45
6	0	6	12	18	24	30	36	42	48	54
7	0	7	14	21	28	35	42	49	56	63
8	0	8	16	24	32	40	48	56	64	72
9	0	9	18	27	36	45	54	63	72	81

Un exemple bien connu est celui des tables d'opérations d'entiers. Une table d'opération est une matrice à 10 lignes et à 10 colonnes (la première ligne et la première colonne étant réservées à 0). A l'intersection de la troisième ligne et de la quatrième colonne, par exemple, se trouve le résultat de l'opération de 2 par 3 (la somme si ce tableau est la table d'addition, le produit s'il s'agit de la multiplication).

Table de multiplication des entiers

L=

température \ objet	O_1	O_2	O_3
T_1	10	9	2
T_2	12	9,7	2,1
T_3	13	10,2	2,1
T_4	13,5	11	2,3

Supposons que nous étudiions la dilatation de trois objets O_1, O_2, O_3 à quatre températures différentes : T_1 (la température initiale), T_2, T_3, T_4. Nous noterons la longueur de ces trois objets, aux différentes températures de l'expérience, dans un tableau à quatre lignes et trois colonnes : 9,7 sera la longueur de l'objet O_2 à la température T_2.

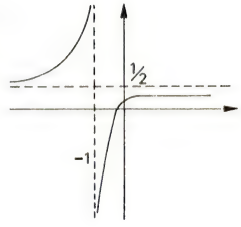
La codification de l'écriture matricielle et l'élaboration du calcul matriciel ont commencé au XIX^e siècle en Angleterre.
C'est par un jeu de résolutions d'équations algébriques et de transformations des inconnues dans ces équations que Cayley imagina d'écrire les coefficients de transformations en tableaux, tandis que Sylvester ◆ avait déjà introduit les matrices.
L'idée de Cayley était celle-ci :
1) A une variable x, associer y par la transformation suivante :

$$y = \frac{ax + b}{cx + d}$$

a, b, c, d étant fixés, pour chaque valeur de x nous pouvons calculer une valeur de y. Par exemple :

$$y = \frac{3x + 2}{6x + 6}$$

◆ A. Cayley, né en 1821, J.J. Sylvester, né en 1814 : mathématiciens anglais qui étudièrent surtout les invariants dans les équations algébriques. Leurs travaux portent également sur la géométrie, l'espace à n dimensions et les matrices.



est une telle fonction qui a pour représentation l'hyperbole ci-contre : Les quatre coefficients a, b, c, d définissent la transformation choisie. Si nous prenons d'autres valeurs de a, b, c, d, nous obtenons une autre transformation dont la représentation est une autre hyperbole. Nous pouvons donc représenter la transformation :

$$y = \frac{ax + b}{cx + d}$$

par le tableau : $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Les lettres ont la même position dans le tableau et dans l'opération. Ce tableau est une matrice.

2) Puis au nombre y, Cayley associa le nombre z par la nouvelle transformation :

$$z = \frac{\mathrm{A}y + \mathrm{B}}{\mathrm{C}y + \mathrm{D}}$$

qu'il représenta par le tableau :

$$\begin{pmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{pmatrix}.$$

Nous pouvons prendre, par exemple :

$$z = \frac{2y + 1}{9y + 2}.$$

Cayley se posa la question : quelle est la transformation qui permet d'associer z à x, et, si elle existe, peut-elle s'écrire, elle aussi, sous forme de matrice? La réponse fut positive. Il appela « produit » des deux matrices précédentes la matrice de résultat. Le calcul matriciel était né (1838).

Revenons sur quelques définitions

Le tableau de nombres s'appelle *matrice.*
Chaque nombre situé dans une case s'appelle *élément de la matrice.* On note souvent a_{ij} l'élément de la i-ième ligne et de la j-ième colonne de la matrice. Dans ce cas, la matrice se note A. Parfois, au nom de la matrice (symbolisée par une lettre latine majuscule) on adjoint un indice formé de deux nombres entiers : $\mathrm{M}_{3,4}$, par exemple. Cela signifie que la matrice M a 3 lignes et 4 colonnes. Si la matrice est composée de nombres réels, nous dirons que *la matrice est réelle* ; si elle est composée de nombres complexes, nous dirons qu'*elle est complexe.*
Deux matrices A et B sont dites égales si A a le même nombre de lignes et de colonnes que B et si tout élément de A est égal à l'élément correspondant (en position) de B, ce qui s'écrit :

$$\mathrm{A} = \mathrm{B} \Leftrightarrow \mathrm{A}_{\mathrm{IJ}} = \mathrm{B}_{\mathrm{IJ}} \text{ et } a_{ij} = b_{ij}.$$

LE CALCUL MATRICIEL ; SON UTILISATION

Nous ne nous intéresserons ici qu'aux matrices réelles (matrices dont les éléments sont des *nombres réels*). Nous avons vu ce qu'est la *multiplication d'une matrice par un nombre :*

Multiplier une matrice M *par un élément x, c'est multiplier chacun de ses termes par x.*

Reprenons l'exemple de la matrice L définie précédemment (longueur des trois objets O_1, O_2, O_3 à quatre températures différentes). Supposons que nous recommencions l'expérience avec trois nouveaux objets U_1, U_2, U_3 ; U_1 étant de la même composition que O_1, mais deux fois et demie plus long que O_1 à la température T_1 (de même pour U_2 et U_3). Notons p_{ij} la longueur de l'objet U_j à la température T_i. Nous aurons : $p_{ij} = 2{,}5\ l_{ij}$.

Soit P la matrice des mesures que nous venons de faire. Nous avons immédiatement : P = 2,5 L.

Cayley avait déjà eu à faire cette opération. Quand il étudia la transformation qui, à x, associe y par l'égalité :

$$y = \frac{ax + b}{cx + d}$$

il l'écrivit, par convention :

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Mais toute autre matrice de la forme :

$$m\mathrm{A} = \begin{pmatrix} ma & mb \\ mc & md \end{pmatrix}$$

représentait la même transformation.

En effet, si, à x, nous associons y de la façon suivante :

$$y = \frac{(ma)\,x + (mb)}{(mc)\,x + (md)}$$

nous avons bien :

$$y = \frac{max + mb}{mcx + md} = \frac{m(ax + b)}{m(cx + d)} = \frac{ax + b}{cx + d} \quad \text{si } m \neq 0$$

Dans cette optique, multiplier une matrice par un nombre non nul, c'était obtenir une nouvelle matrice représentant la même transformation.

Addition de matrices

Revenons à la matrice L représentant la longueur des trois objets O_1, O_2, O_3 à différentes températures au cours d'une expérience.

Supposons maintenant que l'objet O_1 soit formé de deux barres U_1 et V_1 mises bout à bout et que, de la même façon, O_2 soit formé de U_2 et de V_2, et O_3 de U_3 et de V_3. Au lieu d'inscrire la longueur des objets O_1, O_2, O_3, nous pouvons faire deux tableaux : l'un représenterait la longueur des objets U_1, U_2, U_3, mesurée aux quatre températures d'expérience ; l'autre, celle de V_1, V_2, V_3 à ces mêmes températures.

L =

température \ objet	O_1	O_2	O_3
T_1	10	9	2
T_2	12	9,7	2,1
T_3	(13)	10,2	2,1
T_4	13,5	11	2,3

U =

température \ objet	U_1	U_2	U_3
T_1	5	5	0,5
T_2	5,9	5,4	0,5
T_3	(6,3)	5,7	0,5
T_4	6,4	6,1	0,6

V =

température \	V_1	V_2	V_3
T_1	5	4	1,5
T_2	6,1	4,3	1,6
T_3	(6,7)	4,5	1,6
T_4	6,8	4,9	1,7

Il est évident que, si nous voulons connaître la longueur de l'objet O_1 à la température T_3, par exemple, il suffit d'additionner la longueur de U_1 et celle de V_1 à cette même température :

$$13 = 6{,}3 + 6{,}7$$

La matrice L est telle que chacun de ses éléments est la somme de l'élément de U et de celui de V placés dans la même position que lui dans la matrice correspondante.

> $l_{ij} = u_{ij} + v_{ij}$ pour tout i et j.
> L est la somme de U et de V.

Multiplication des matrices entre elles

C'est Cayley lui-même, auteur du calcul matriciel, qui a défini leur multiplication. L'idée de Cayley était la suivante : ayant rapporté, dans un premier tableau T_1, une transformation qui, à x, associait y suivant le processus

$$y = \frac{ax + b}{cx + d} \qquad T_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

puis, dans un second tableau

$$T_2 = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

une seconde transformation qui, à y, associait z suivant le processus

$$z = \frac{Ay + B}{Cy + D}$$

existait-il un tableau T_3 à deux lignes et deux colonnes donnant la transformation permettant de passer directement de x à z? Il suffisait, pour cela, d'effectuer le calcul donnant z en fonction de x de la façon suivante :

$$z = \frac{Ay + B}{Cy + D} = \frac{A\left(\frac{ax + b}{cx + d}\right) + B}{C\left(\frac{ax + b}{cx + d}\right) + D}$$

$$= \frac{Aax + Ab + Bcx + Bd}{Cax + Cb + Dcx + Db}$$

$$= \frac{(Aa + Bc)\,x + Ab + Bd}{(Ca + Dc)\,x + Cb + Dd}$$

Si nous reprenons l'exemple précédent (page 90) nous obtenons :

$$y = \frac{3x + 2}{6x + 6} \qquad T_1 = \begin{pmatrix} 3 & 2 \\ 6 & 6 \end{pmatrix}$$

$$z = \frac{2y + 1}{9y + 2} \qquad T_2 = \begin{pmatrix} 2 & 1 \\ 9 & 2 \end{pmatrix}$$

$$z = \frac{2\left(\frac{3x + 2}{6x + 6}\right) + 1}{9\left(\frac{3x + 2}{6x + 6}\right) + 2} = \frac{2(3x + 2) + 1(6x + 6)}{9(3x + 2) + 2(6x + 6)} = \frac{12x + 10}{39x + 30}$$

$$\begin{pmatrix} 2 & 1 \\ 9 & 2 \end{pmatrix} \times \begin{pmatrix} 3 & 2 \\ 6 & 6 \end{pmatrix} = \begin{pmatrix} 12 & 10 \\ 39 & 30 \end{pmatrix}$$

La transformation qui permet de passer de x à z est du même type que les précédentes. Ses coefficients sont : $Aa + Bc$, $Ab + Bd$, $Ca + Dc$, $Cb + Dd$ que nous pouvons écrire dans le tableau :

$$T_3 = \begin{pmatrix} Aa + Bc & Ab + Bd \\ Ca + Dc & Cb + Dd \end{pmatrix}$$

T_3 est le produit de T_2 par T_1 :

$$T_3 = T_2 \times T_1.$$

Il existe beaucoup d'autres utilisations du calcul matriciel.
La première utilisation vraiment spectaculaire des matrices fut celle de Heisenberg, en 1925, pour rendre compte des *phénomènes quantiques* dans la structure atomique.

Nous avions annoncé l'étude de la géométrie classique par l'intermédiaire de l'algèbre linéaire. A présent, nous sommes seulement arrivés à localiser les points d'un espace par l'intermédiaire des vecteurs et à définir le parallélisme. La notion de distance et d'angle n'est pas encore définie.

Qu'est-ce que la distance de deux points A et B? C'est une application qui à tout couple de points A et B associe un nombre réel positif d (A, B), qui jouit des propriétés suivantes :

$\forall$ (A, B) (1) d (A, B) $\geqslant 0$.

(2) d (A, B) $= 0 \iff$ A et B sont confondus.

(3) d (A, B) $= d$ (B, A).

et (4) d (A, C) $\leqslant d$ (A, B) $+ d$ (B, C), appelé inégalité triangulaire.

Au lieu de définir directement la distance sur l'ensemble des points, nous allons définir la norme des vecteurs joignant deux points (sorte de « mesure » de ce vecteur), application de l'ensemble des vecteurs dans l'ensemble des nombres réels $\mathbb{R}$, et nous poserons :

$$d\,(A, B) = N\,(OB - OA) = N\,(AB).$$

Pour que la norme d'un vecteur permette de définir la distance de deux points extrémités de ce vecteur, elle devra vérifier les propriétés suivantes :

$\forall\ V \in E$ (1) $N\,(V) \geqslant 0$.

(2) $N\,(V) = O \iff V = O_E$.

(3) $N\,(V) = N\,(-V)$.

(4) $N\,(V_1 + V_2) \leqslant N\,(V_1) + N\,(V_2)$.

Elle vérifiera aussi la propriété :

(5) $N\,(\alpha\, V) = |\,\alpha\,| \,.\, N\,(V)$.

La propriété (3) est un cas particulier de (5).

En fait, une norme est très souvent définie à partir d'un produit scalaire ◆. Cette nouvelle application : « produit scalaire », apparemment plus compliquée que l'application : « norme », est en fait beaucoup plus riche.

◆ Voir aussi ce mot au dictionnaire.

Le produit scalaire de deux vecteurs V_1 et V_2 est un nombre réel, noté $< V_1, V_2 >$ ◆ ; l'application qui, à chaque couple (V_1, V_2) de vecteurs de E, fait correspondre leur produit scalaire est une application, que l'on peut noter $<, >$, de E $\times$ E dans $\mathbb{R}$, bilinéaire, symétrique et positive, c'est-à-dire qu'elle vérifie les trois propriétés suivantes :

◆ On le note aussi, souvent, $V_1 \cdot V_2$.

$$<, > : E \times E \longrightarrow \mathbb{R}$$
$$(V_1, V_2) \longmapsto < V_1, V_2 > = r \in \mathbb{R}$$

(1) — Si V_1 est fixé et si V_2 varie dans E,
alors $< V_1, V >$ est linéaire par rapport à V.

— Si V_2 est fixé et si V_1 varie dans E,

alors $< V, V_2 >$ est linéaire par rapport à V.

(2) Symétrique : $\forall (V_1, V_2) \in E \times E$,

$$< V_1, V_2 > = < V_2, V_1 >.$$

(3) Positive : $\forall (V_1, V_2) \in E \times E$,

$$< V_1, V_2 > \geqslant 0.$$

Un produit scalaire est dit *défini positif* s'il vérifie en plus la propriété :

(4) $\forall V \in E, \quad < V, V > = 0 \iff V = O_E.$

Un produit scalaire défini positif permet de définir une norme. On pose, dans ce cas, par convention :

$$N(V) = \sqrt{< V, V >}$$

On peut montrer que l'expression N(V) ainsi définie vérifie bien les cinq propriétés de la norme, donc permet de définir une distance dans l'espace affine correspondant.

Mais le produit scalaire permet alors de définir aussi l'angle de deux vecteurs. Dans ce cas, on pose :

$$\cos (V_1, V_2) = \frac{< V_1, V_2 >}{N(V_1) . N(V_2)}$$

qui se lit : « Le cosinus de l'angle formé par les vecteurs V_1 et V_2 est le rapport du produit scalaire de ces deux vecteurs divisé par le produit de leurs normes. » Une étude approfondie de cette définition permettrait de montrer que l'angle ainsi mesuré peut coïncider avec l'angle défini dans la géométrie euclidienne.

L'algèbre linéaire nous a fait découvrir l'importance de la structure d'espace vectoriel (structure d'ensembles aussi importants que celui des fonctions réelles à valeurs réelles, par exemple). Elle nous a également montré la diversité des domaines d'utilisation du calcul matriciel. Ces raisons sont suffisantes pour en faire un domaine particulier de l'algèbre.

Danielle Fèvre.

ALGÈBRE (Structure d') ◆

◆ La définition donnée ci-dessous faisant intervenir un grand nombre de notions d'algèbre, il est conseillé au lecteur non familiarisé avec ces notions de lire l'article « l'Algèbre » en entier et Vectoriel (Espace).

Un ensemble E a structure d'algèbre sur le corps K (on dit aussi K-algèbre) s'il est muni de trois lois de composition : deux lois de composition interne (l'une notée $+$ et l'autre $\times$) et une loi de composition externe (notée $\bullet$) satisfaisant aux conditions suivantes :

1) E muni de la loi $+$ et de la loi $\bullet$ est un espace vectoriel sur K.

2) La loi $\times$ est distributive par rapport à la loi $+$.

3) Pour tout couple (λ, μ) d'éléments de K et tout couple (x, y) d'éléments de E, on a : $(\lambda.x) \times (\lambda.y) = (\lambda \times \mu).(x \times y)$.
Si la loi $\times$ est associative (donc si E muni de $+$ et de $\times$ est un anneau), l'algèbre est dite associative ; si la loi $\times$ est commutative, l'algèbre est dite commutative.

ALGÈBRE DE BOOLE pages 329, 417.

George Boole (1815-1864), en publiant « Analyse mathématique de la logique » (1847) et « Étude des lois de la pensée » (1854), a donné une armature rigoureuse, de type mathématique, à la logique. La démarche consiste à repérer certaines fonctions caractéristiques dans notre langage ordinaire, de façon à pouvoir les traduire dans les termes d'une algèbre symbolique et à identifier dans celles-ci des lois de calcul qui deviendront la forme mathématique des lois de la pensée. Pour donner un exemple, deux expressions équivalentes se traduiront par une propriété de commutativité : $xy = yx$.
Le fait que la base de la logique soit la proposition, c'est-à-dire un énoncé qui est soit vrai, soit faux, se traduira de même par le fait que, dans l'algèbre mise au point par Boole, il n'y a que deux valeurs : 1 (pour *vrai*) et 0 (pour *faux*).
Une algèbre de Boole◆ est un ensemble E muni de deux lois de composition interne $\top$ et $\bot$ et d'une application f de E dans lui-même qui à chaque élément x de E fait correspondre un élément de E noté $\bar{x}$◆, vérifiant les propriétés suivantes :

◆ Il est à noter qu'une algèbre de Boole n'est pas une algèbre au sens de la définition de la structure d'algèbre.

◆ On lit « x barre ».

1) Chacune des deux lois $\top$ et $\bot$ est commutative et associative :

$$a \top b = b \top a \qquad a \top (b \top c) = (a \top b) \top c$$
$$a \bot b = b \bot a \qquad a \bot (b \bot c) = (a \bot b) \bot c.$$

2) Chaque élément de E est idempotent pour les deux lois :

$$a \top a = a \quad \text{et} \quad a \bot a = a.$$

3) Chacune des deux lois est distributive par rapport à l'autre :

$$a \top (b \bot c) = (a \top b) \bot (a \top c)$$
$$a \bot (b \top c) = (a \bot b) \top (a \bot c).$$

4) La loi $\top$ admet un élément neutre $e_\top$ et la loi $\bot$ un élément neutre $e_\bot$.

5) Pour tout élément a de E :

$$a \bot \bar{a} = e_\top \quad \text{et} \quad a \top \bar{a} = e_\bot.$$

6) Quel que soit a de E, si $\bar{a}$ est l'élément qui lui correspond par f, l'élément correspondant à $\bar{a}$ par f est a : $\bar{\bar{a}} = a$.

7) Quels que soient a et b de E, l'élément correspondant à $a \top b$ par f est égal à $\bar{a} \bot \bar{b}$:

$$\overline{(a \top b)} = \bar{a} \bot \bar{b}.$$

L'élément correspondant à $a \perp b$ par f est $\bar{a} \top \bar{b}$: $\overline{(a \perp b)} = \bar{a} \top \bar{b}$.

Exemple

L'ensemble $\mathfrak{P}(E)$ des parties d'un ensemble E, muni de la loi intersection (jouant le rôle de $\top$), de la réunion (jouant le rôle de $\perp$), de l'application qui à un élément A de $\mathfrak{P}(E)$ fait correspondre son complémentaire dans E est une algèbre de Boole ◆. Les propositions logiques ont une structure analogue : « et » joue le rôle de $\top$, « ou » celui de $\perp$, l'application $x \longmapsto \bar{x}$ est ici la négation (le signe $=$, quand il s'agit de propositions logiques, est remplacé par le signe $\Longleftrightarrow$).

◆ Réciproquement, pour toute algèbre de Boole E, il existe un ensemble F tel que E soit isomorphe à une partie de $\mathfrak{P}(F)$.

Anneau de Boole : un ensemble E, qui a structure d'anneau pour les deux lois de composition interne $+$ et $\times$, est un anneau de Boole si, quel que soit l'élément x de E, $x \times x = x$ (tout élément x de E est idempotent pour la loi $\times$).
L'ensemble $\mathfrak{P}(E)$ des parties d'un ensemble E, muni des deux lois de composition interne, différence symétrique et intersection, est un anneau de Boole.

Treillis de Boole : voir Treillis.

ALGÉBRIQUE pages 373 et 374.

Équation algébrique : voir Équations.

Mesure algébrique
Soient A et B deux points situés sur un axe Ox. On appelle mesure algébrique du segment AB et on note $\overline{AB}$ le nombre réel (positif ou négatif) égal à la différence entre l'abscisse de B et celle de A : $\overline{AB} = x_B - x_A$.

Exemples

$$\left.\begin{array}{l}\text{Abscisse de B} = 3\\ \text{Abscisse de A} = 1\end{array}\right\} \overline{AB} = 3 - 1 = 2$$

$$\left.\begin{array}{l}\text{Abscisse de A} = 5\\ \text{Abscisse de B} = 2\end{array}\right\} \overline{AB} = 2 - 5 = -3$$

La mesure algébrique de AB et celle de BA sont opposées : $\overline{AB} = -\overline{BA}$.
Si A, B, C sont trois points d'un même axe, les mesures algébriques de AB, de BC et de CA sont liées par la relation :

$$\overline{AB} + \overline{BC} + \overline{CA} = 0.$$

Cette relation est appelée relation de Chasles.

Nombre algébrique : voir l'article « les Nombres », page 373.

ALGORITHME pages 179 et 180.

En arithmétique, en informatique, etc. : suite d'opérations permettant de résoudre un problème particulier.

Exemple : algorithme d'Euclide permettant de trouver le plus grand commun diviseur (p.g.c.d.) de deux nombres a et b. On commence par diviser a par b,

Figure 1

6' 52' 32' 45 |

Figure 2

6 52 32 45 | 2
2

Figure 3

6 52 32 45 | 2
2 52 |
45 × 5 = 225

Figure 4

6 52 32 45 | 25
2 52 | 45 × 5 = 225
2 25 | 505 × 5 = 2525
27 32

Figure 5

6 52 32 45 | 2554
2 52 | 45 × 5 = 225
2 25 | 505 × 5 = 2525
27 32 | 5104 × 4 = 26435
25 25
2 07 45 |

puis *b* par le reste de cette première opération, puis ce dernier reste par le second, etc. On arrive, après un certain nombre d'opérations, à un reste nul. Le reste précédent est le p.g.c.d.
Calculs faits pour trouver le p.g.c.d. de 75 et de 45 :

75 : 45 = 1, reste 30 ;
45 : 30 = 1, reste 15 ;
30 : 15 = 2, reste 0.

Le p.g.c.d. de 75 et de 45 est donc 15.

Algorithme de l'extraction de la racine carrée
Supposons vouloir extraire la racine carrée du nombre 6 523 245.

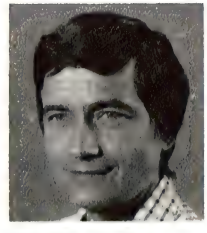
On dispose une opération analogue à une division.
On sépare le nombre en tranches de deux chiffres à partir de la droite (fig. 1).
On cherche le plus grand nombre dont le carré soit inférieur à la tranche de gauche (ici : 6) ; ce nombre est 2.
On inscrit ce nombre à droite et on soustrait son carré de la tranche de gauche (ici : 6 — 4 = 2) (fig. 2).
On place à côté du résultat obtenu la tranche suivante de deux chiffres (52).
On multiplie le nombre situé au-dessus du trait horizontal par 2 et on l'inscrit en dessous du trait horizontal (ici : 4).
On cherche quel est le plus grand chiffre x que l'on doit inscrire à la suite de 4 pour que $4x \times x$ soit plus petit que 252 ; on trouve $45 \times 5 = 225$, qui est bien inférieur à 252 (alors que $46 \times 6 = 276$ est supérieur à 252) (fig. 3).
On inscrit 5 à côté du 2 (fig. 4). On soustrait 225 de 252 : on trouve 27. On écrit à côté de 27 la tranche suivante de deux chiffres : 32. On multiplie 25 par 2 : 50 et on cherche le plus grand chiffre x que l'on doit inscrire à la suite de 50 pour que $50x \times x$ soit inférieur à 2 732 : c'est 5.
On l'écrit à la suite de 25 : 255 (fig. 5).
On retranche 2 525 de 2 732 et on obtient 207.
On inscrit à côté de 207 la tranche suivante de deux chiffres : 45. On double 255 = 510 et on cherche quel est le chiffre x qu'il faut mettre à la suite de 510 pour que $510x \times x$ soit inférieur à 20 745 : c'est 4. On l'inscrit après 255 : 2 554.
Il n'y a plus de tranche de deux chiffres à utiliser ; on s'arrête donc.
La racine carrée à une unité près de 6 523 245 est 2 554.

ANALYSE : voir l'article pages suivantes et les pages 157, 327 à 330, 388.

L'analyse

par Christiane de Bary

Nous nous proposons dans les quelques chapitres qui vont suivre d'exposer quelques points de la théorie des fonctions d'une *variable réelle à valeurs réelles* et *des suites de nombres réels*. Nous indiquerons rapidement comment ces théories se généralisent aux fonctions de plusieurs variables réelles. Cette étude reste à la base de l'analyse mathématique moderne.



Christiane de Bary
Née à Bordeaux le 7 décembre 1931. Agrégée de l'Université, maître-assistant à la Faculté des Sciences de Metz. Poursuit actuellement des recherches en documentation automatique à l'Institut universitaire de calcul automatique de Nancy.

Que signifie l'expression *fonction d'une variable réelle à valeurs réelles?*
Prenons un exemple : « Je veux étudier la longueur d'une barre d'un métal en fonction de la température. » Je mesure que, si la température est de :

10 °C, la longueur de la barre est 1 m ;

20 °C, la longueur de la barre est 1,05 m ;

30 °C, la longueur de la barre est 1,10 m ;

40 °C, la longueur de la barre est 1,15 m ;

50 °C, la longueur de la barre est 1,20 m.

Je m'aperçois donc que, si la température augmente, la longueur varie : la longueur de la barre est *fonction* de la température.
Je répète les expériences, et je vois que la longueur y de la barre, mesurée en mètres, est liée à la température x mesurée en °C par une relation qui peut s'écrire :

$$y - y_0 = 0{,}05 \times (x - x_0)$$

où x_0 vaut 10 °C et y_0 1 m, longueur de la barre à 10 °C.
Mais, dès que la température est inférieure à — 18 °C ou supérieure à 340 °C, je me rends compte que cette formule n'est plus valable. Cette correspondance entre le nombre réel qui représente la température et celui qui représente la longueur est une fonction définie

seulement pour x compris entre — 18 °C et 340 °C. Les mathématiciens la représentent ainsi :

$$f : [-18, 340] \longrightarrow \mathbb{R}$$
$$x \longmapsto y = y_0 + 0{,}05\,(x - x_0)$$
$$y = 1 + (0{,}05)\,(x - 10)$$

D'une manière plus générale, une fonction d'une variable réelle à valeurs réelles est une règle qui à tout réel d'une partie D de l'ensemble des nombres réels $\mathbb{R}$ (et pas nécessairement $\mathbb{R}$ tout entier) associe un autre réel. Pour une suite, l'ensemble D est l'ensemble des nombres entiers positifs. Ainsi, une suite de nombres réels est une règle qui à chaque entier n positif associe un nombre réel.
Pour faire une étude cohérente de ces fonctions, nous nous rendons compte que nous sommes amenés à préciser les propriétés de l'ensemble des réels. Nous passerons ensuite à l'étude des suites, puis à celle des fonctions.

L'ENSEMBLE DES NOMBRES RÉELS $\mathbb{R}$

Nous allons donner un certain nombre de propriétés de l'ensemble des nombres réels $\mathbb{R}$, 16 au total. Toutes les propriétés de $\mathbb{R}$, aussi bien les règles de calcul bien connues que les propriétés les plus fines, peuvent s'en déduire en toute rigueur ◆. Cette exigence n'empêche pas de nous *représenter* $\mathbb{R}$ en associant à tout réel x le point M d'abscisse $\overline{OM} = x$ d'un axe.
Cette *représentation géométrique* peut aider à découvrir un raisonnement. Mais une « évidence » sur la représentation géométrique ne remplace pas pour un mathématicien une démonstration.

◆ Cela revient, en fait, à faire une construction axiomatique de l'ensemble des réels. C'est la méthode adoptée par Dieudonné dans son ouvrage *les Fondements de l'analyse moderne*.

PROPRIÉTÉS FONDAMENTALES DE $\mathbb{R}$

Les quinze premières propriétés sont relatives aux opérations, addition, multiplication, et à la relation d'inégalité. La seizième est particulière à l'ensemble des nombres réels.

1. — *L'addition est commutative.*
Quels que soient x, y $x + y = y + x$.

2. — *L'addition est associative.*
Quels que soient x, y, z $x + (y + z) = (x + y) + z$.
On pose alors par définition :
$$x + y + z = x + (y + z) = (x + y) + z.$$

3. — Il existe un *élément noté* 0 tel que,
pour tout x, $x + 0 = x$.
0 est le seul élément jouissant de cette propriété : il est appelé élément neutre pour l'addition.

4. — Quel que soit a, il existe un élément x et un seul tel que $x + a = 0$.
Il est appelé *opposé de a* et noté $-a$ ◆.

5. — *La multiplication est commutative.*

Quels que soient x, y $x \times y = y \times x$.

6. — *La multiplication est associative.*

Quels que soient x, y, z $x \times (y \times z) = (x \times y) \times z$.

On pose alors par définition

$$x \times y \times z = x \times (y \times z) = (x \times y) \times z.$$

7. — *La multiplication est distributive par rapport à l'addition.*

Quels que soient x, y, z $x \times (y + z) = (x \times y) + (x \times z)$.

8. — Il existe un *élément noté* 1 tel que, pour tout x, $x \times 1 = x$.

1 est le seul élément jouissant de cette propriété : il est appelé élément neutre pour la multiplication.

9. — Quel que soit $a \neq 0$, il existe un élément x et un seul tel que $x \times a = 1$
x est appelé *inverse de a* et noté $\frac{1}{a}$ ◆.

10. — Quels que soient x et y on a $x \leqslant y$ ou $y \leqslant x$.

11. — Quels que soient x, y, z si $x \leqslant y$ et $y \leqslant z$
alors $x \leqslant z$.

12. — Quels que soient $x, y,$ si $x \leqslant y$ et $y \leqslant x$ alors $x = y$.

13. — Quels que soient $x, y, z,$ si $x \leqslant y$ alors $x + z \leqslant y + z$.

14. — Quels que soient x, y si $0 \leqslant x$ et $0 \leqslant y$ alors $0 \leqslant xy$ ◆.

15. — Quels que soient $x > 0$ $y \geqslant 0$ il existe un entier n tel que $y \leqslant nx$.

Remarquons que cette propriété peut encore s'énoncer :
Il existe une infinité d'entiers n tels que $y \leqslant nx$ ◆.

16. — La seizième propriété fondamentale est la seule qui soit particulière à l'ensemble $\mathbb{R}$ des nombres réels qui ne soit pas vraie dans l'ensemble $\mathbb{Q}$ des nombres rationnels. Avant de l'énoncer, nous devons donner quelques précisions.

◆ Ces 4 propriétés donnent à $\mathbb{R}$ une structure de groupe commutatif pour l'addition.
Voir l'article « l'Algèbre », page 28.

◆ Les 9 propriétés énoncées donnent à $\mathbb{R}$ une structure de corps commutatif.
Voir l'article « l'Algèbre », page 53.

◆ La relation « $\leqslant$ » jouissant des propriétés 10 à 14 est une relation *d'ordre total*. Voir l'article « les Ensembles ».

◆ On peut déduire des 15 premières propriétés fondamentales toutes les règles de calcul que l'on connaît depuis l'enfance, par exemple :
$x \times 0 = 0$
si $x + a = y + a$
alors $x = y$
si $x \times a = y \times a$
alors $x = y$ si $a \neq 0$
etc.

◆ Lire également Intervalle.

Intervalles ◆, *intervalles emboîtés*

Dans $\mathbb{R}$, on appelle *intervalle fermé d'origine a, d'extrémité b*, l'ensemble des nombres réels x tels que $a \leqslant x \leqslant b$.

$$[a, b] = \{ x \in \mathbb{R} \mid a \leqslant x \leqslant b \}$$

(lire : l'intervalle a, b est égal à l'ensemble des x appartenant à $\mathbb{R}$ tels que $a \leqslant x \leqslant b$).

La longueur de l'intervalle $[a, b]$ est égale au nombre réel $b - a$.

Exemples

Ainsi $I_1 = [-1, +1] = \{ x \in \mathbb{R} \mid -1 \leqslant x \leqslant +1 \}$

$I_2 = \left[-\frac{1}{2}, +\frac{1}{2}\right] = \left\{ x \in \mathbb{R} \mid -\frac{1}{2} \leqslant x \leqslant +\frac{1}{2} \right\}$

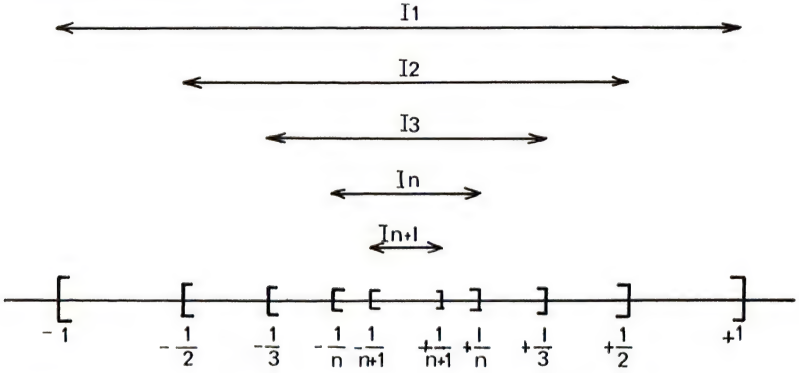
$I_3 = \left[-\frac{1}{3}, +\frac{1}{3}\right] = \left\{ x \in \mathbb{R} \mid -\frac{1}{3} \leqslant x \leqslant +\frac{1}{3} \right\}$

.....

sont des intervalles fermés de $\mathbb{R}$.

Et $I_1, I_2, I_3, \ldots, I_n = \left[-\frac{1}{n}, +\frac{1}{n}\right], I_{n+1} = \left[-\frac{1}{n+1}, +\frac{1}{n+1}\right] \ldots$

forment une *suite d'intervalles fermés emboîtés.*



Cette suite d'intervalles est dite emboîtée car, pour tout entier n :

$I_{n+1} \subset I_n$.

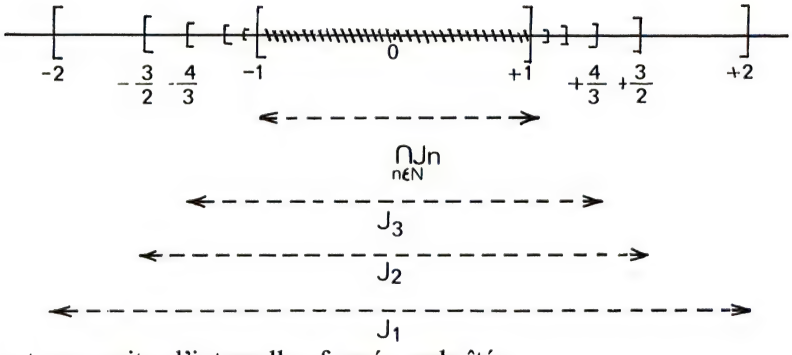
De même, la suite d'intervalles fermés

$J_1 = \left[-\frac{2}{1}, +\frac{2}{1}\right]$ $\qquad$ $J_3 = \left[-\frac{4}{3}, +\frac{4}{3}\right]$

.....

$J_2 = \left[-\frac{3}{2}, +\frac{3}{2}\right]$ $\qquad$ $J_n = \left[-\frac{n+1}{n}, +\frac{n+1}{n}\right]$

.....



est une suite d'intervalles fermés emboîtés.

Dans le premier exemple, il existe un point 0 et un seul commun à tous les intervalles I_n. On écrit :

$$\bigcap_{n\in N} I_n = \{ 0 \}$$

(on lit : l'intersection de la suite d'intervalles I_n est réduite au point 0).

Dans le deuxième exemple, tout réel x tel que $-1 \leqslant x \leqslant +1$ appartient à tous les intervalles J_n et il n'y a pas d'autre réel appartenant à tous ces intervalles J_n.

$$\bigcap_{n\in N} J_n = [-1, +1]$$

Peut-il exister une suite d'intervalles fermés emboîtés pour laquelle il n'y aurait aucun point commun à tous les intervalles I_n?

16. — La propriété fondamentale 16, dite *propriété des intervalles fermés emboîtés* ◆, permet d'affirmer que non. Elle s'énonce :
Propriété 16 :
Toute suite d'intervalles fermés I_n *emboîtés a une intersection non vide :*

$$\bigcap_{n\in N} I_n \neq \varnothing$$

(on lit : l'ensemble de points appartenant à tous les intervalles I_n est non vide, $\neq$: différent de, $\varnothing$: ensemble vide).

◆ On peut considérer que cette propriété fondamentale de $\mathbb{R}$, qui n'est pas vraie dans $\mathbb{Q}$, répond à l'exigence intuitive (article « les Nombres », page 385). « $\mathbb{R}$ est une extension de $\mathbb{Q}$ dans laquelle il n'y a plus de trous. »

Dans le cas particulier d'une suite d'intervalles fermés emboîtés

$$I_n = [a_n, b_n]$$

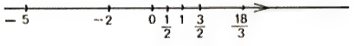
dont la longueur des intervalles $b_n - a_n$ *tend vers* 0 ◆, cette intersection est réduite à un point.

◆ Voir le sens exact de cette expression page 109.

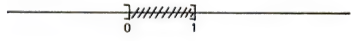
ENSEMBLE MAJORÉ, MINORÉ, BORNÉ

Considérons les ensembles de nombres réels suivants:

$$F = \left\{ 1, \frac{3}{2}, -5, \frac{1}{2}, -2, \frac{18}{3} \right\}$$



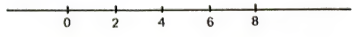
$$G = \{ x \in \mathbb{R} \mid 0 < x \leqslant 1 \} \blacklozenge$$



◆ Lire : ensemble des x appartenant à l'ensemble des réels tels qu'ils soient compris entre 0 et 1.

$$H = \{ x \in \mathbb{R} \mid x < 0 \}$$

$$K = \{ x \in \mathbb{R} \mid x = 2n, n \in \mathbb{N} \}$$



K n'est autre que l'ensemble des entiers pairs positifs ou nuls.

L'ensemble F contient un élément $\frac{18}{3}$ qui est supérieur ou égal à tout élément de F. On dit que $\frac{18}{3}$ est le « plus grand élément de F ».

L'ensemble G contient aussi un plus grand élément 1.

L'ensemble H, au contraire, ne contient pas de plus grand élément, pas plus que l'ensemble K qui est formé des entiers pairs, positifs. Mais l'ensemble H est majoré par 0, alors que l'ensemble K n'est pas majoré.

Ensemble majoré

Un ensemble $E \subset \mathbb{R}$ (E inclus dans $\mathbb{R}$), non vide ◆, est majoré s'il existe un réel A tel que

◆ C'est-à-dire contenant au moins 1 élément.

$$x \in E \longrightarrow x \leqslant A$$

(lire : x appartenant à E implique x inférieur ou égal à A).

A est appelé *majorant de l'ensemble* E. On dit encore que E *est majoré par* A.

Un ensemble formé d'un nombre fini d'éléments a toujours un plus grand élément et, par suite, est majoré par ce plus grand élément (ensemble F).

Un ensemble infini peut :

1) avoir un plus grand élément ; il admet alors ce plus grand élément comme majorant (ensemble G) ;

2) ne pas avoir de plus grand élément tout en étant majoré (ensemble H) ;

3) ne pas être majoré (ensemble K).

Ensemble minoré

De même, l'ensemble F et l'ensemble K admettent un *plus petit élément*. L'ensemble G n'admet pas de plus petit élément, mais il *est minoré par* 0. L'ensemble H n'est pas minoré.
Un ensemble $E \subset \mathbb{R}$ non vide est minoré s'il existe un réel A tel que

$$x \in E \Longrightarrow x \geqslant A$$

Les propriétés vues pour les majorants (cas d'un ensemble fini et d'un ensemble infini) s'énonceraient de façon identique pour les minorants.

Plus petit majorant, plus grand minorant

Revenons alors à un ensemble E non vide de $\mathbb{R}$ que nous supposons majoré. Désignons par E* l'ensemble des majorants de E.
E* est non vide puisque E est majoré.
E* est infini ; en effet, si A est un majorant de E, tout nombre $A' > A$ est à fortiori un majorant de E.
E* est donc un ensemble infini minoré ; il peut avoir un plus petit élément ou ne pas en avoir.
Nous pourrions démontrer que E* admet un plus petit élément. Cette propriété, connue sous le nom de *théorème de Bolzano-Weierstrass*, s'énonce ◆:

Tout ensemble $E \subset \mathbb{R}$, *non vide, majoré, admet un majorant inférieur à tous les autres.*

De même que *tout ensemble* $E \subset \mathbb{R}$, *non vide, minoré, admet un minorant supérieur à tous les autres.*

◆ Le théorème de Bolzano-Weierstrass n'est pas vrai dans $\mathbb{Q}$ (ensemble des rationnels). Nous aurions pu le choisir comme 16e propriété fondamentale, à la place de la propriété des suites d'intervalles emboîtés. Il se démontre d'ailleurs par l'intermédiaire de cette 16e propriété.

Le plus petit majorant d'un ensemble $E \subset \mathbb{R}$ est appelé *borne supérieure de* E, noté : sup E ou $\sup_{x \in E} x$.

Le plus grand minorant d'un ensemble $E \subset \mathbb{R}$ est appelé *borne inférieure de* E, noté : inf E ou $\inf_{x \in E} x$.

Reprenons les exemples précédents :

$$\sup F = \frac{18}{3} \qquad \inf F = -5$$

$$\sup G = 1 \qquad \inf G = 0$$

$$\sup H = 0$$

$$\inf K = 0$$

Si un ensemble admet un plus grand élément, celui-ci est bien évidemment son plus petit majorant (sa borne supérieure).
Un ensemble $E \subset \mathbb{R}$, non vide, est borné s'il est à la fois majoré et minoré.
Le théorème de Bolzano-Weierstrass est fondamental en analyse.

SUITE DE NOMBRES RÉELS

Se donner une suite de réels, c'est se donner une règle qui à tout nombre entier $n > 0$ permet d'associer un nombre réel ◆.
Le nombre réel associé à l'entier 1 est appelé le premier terme de la suite et noté u_1.
Le nombre réel associé à l'entier 2 est appelé le deuxième terme de la suite et noté u_2.
Le nombre réel associé à l'entier 3 est appelé le troisième terme de la suite et noté u_3.
Le nombre réel associé à l'entier n est appelé le n-ième terme de la suite et noté u_n.

◆ Si l'on veut définir avec rigueur la notion de suites de nombres réels, on doit dire qu'une suite de nombres réels est une application de l'ensemble des nombres entiers naturels $\mathbb{N}$ dans l'ensemble des nombres réels.

Ainsi, par exemple, nous pouvons envisager la suite dont le premier terme est 1, le deuxième terme $\frac{1}{2}$, le troisième terme $\frac{1}{3}$, le quatrième terme $\frac{1}{4}$, le n-ième terme $\frac{1}{n}$.

Nous pouvons aussi envisager une suite dont le premier terme u_1 est égal à 6, le deuxième terme u_2 à 999 et, pourquoi pas, le centième terme u_{100} égal à 0,5, et indéfiniment.

$$\begin{aligned}
&1^{er} \text{ terme} && u_1 = 6\\
&2^{e} \text{ terme} && u_2 = 999\\
&3^{e} \text{ terme} && u_3 = 10{,}5\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&100^{e} \text{ terme} && u_{100} = 0{,}5\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

Différentes façons de se donner une suite

Le plus souvent, quand nous étudions une suite, nous nous donnons l'expression de son n-ième terme u_n en fonction de n et, pour désigner une telle suite, nous parlons de la suite $u_n = f(n)$. Ainsi, dans le premier exemple, nous pouvons définir la suite par $u_n = f(n) = \frac{1}{n}$.
Cette suite est appelée *suite harmonique*.
Mais une suite peut aussi être définie par la donnée de son premier

terme u_1 et par une relation qui permet de calculer un terme en fonction de celui qui le précède. Par exemple, nous pouvons envisager la suite définie par

$$u_1 = \frac{1}{3} \qquad u_{n+1} = \frac{1}{2}(u_n^2 + 1).$$

Si, dans la relation $u_{n+1} = \frac{1}{2}(u_n^2 + 1)$, nous faisons $u_1 = \frac{1}{3}$, nous obtenons

$$u_2 = \frac{1}{2}(u_1^2 + 1) = \frac{1}{2}\left[\left(\frac{1}{3}\right)^2 + 1\right] = \frac{5}{9}.$$

Pour $n = 2$, nous obtenons

$$u_3 = \frac{1}{2}(u_2^2 + 1) = \frac{1}{2}\left[\left(\frac{5}{9}\right)^2 + 1\right] = \frac{53}{81}, \text{ etc.}$$

Suite somme et suite produit

On peut aussi, étant donné deux suites quelconques u_n et v_n, envisager la suite de n-ième terme $u_n + v_n$ ou la suite de n-ième terme $u_n \times v_n$. Ces suites sont respectivement appelées :

suite somme des suites u_n et v_n

suite produit des suites u_n et v_n

Prenons par exemple les suites :

$$u_n = \frac{1}{n} \qquad v_n = 2n.$$

La suite $u_n + v_n$ a pour n-ième terme $\frac{1}{n} + 2n$, la suite $u_n \times v_n$: $2n \times \frac{1}{n} = 2$.

Tous les termes de la suite $u_n \times v_n$ sont identiques. Une telle suite dont tous les termes sont identiques est dite *stationnaire*. En effet, $2n \times \frac{1}{n} = \frac{2n}{n} = 2$.

On définit de même :

Le produit d'une suite u_n par un réel a. Ainsi, la suite $v_n = 2n$ n'est autre que le produit de la suite de n-ième terme n par le nombre réel 2 ◆.

◆ $v_n = 2n$ peut aussi être considéré comme le produit de la suite de n-ième terme n par la suite stationnaire dont tous les termes sont égaux à 2.

La différence de deux suites u_n *et* v_n et, si tous les termes de la suite sont différents de 0, *la suite inverse de la suite* u_n ou *la suite quotient d'une suite* v_n *par* u_n.

$$u_n - v_n = \frac{1}{n} - 2n \qquad \frac{v_n}{u_n} = \frac{2n}{\frac{1}{n}} = 2n^2.$$

Limite d'une suite (finie et infinie). Convergence et divergence

Le plus important dans l'étude d'une suite est, en général, de déterminer ce que devient u_n pour n très grand. Considérons, ainsi, la suite $u_n = \frac{3}{2}n$. La propriété $u_n \geqslant 10\,000$ n'est pas vraie pour tout entier n. Toutefois, si $n \geqslant 8\,000$, alors $u_n \geqslant 10\,000$. Une telle propriété d'une suite u_n, vraie à partir d'un certain indice, est dite « *vraie pour n assez grand* ».

Considérons maintenant la suite $u_n = \frac{n-1}{n} = 1 - \frac{1}{n}$.

Choisissons un réel ε positif. La propriété $|u_n - 1| \leqslant \varepsilon$ est vraie pour n assez grand puisque $|u_n - 1| = \frac{1}{n}$.

Cela reste vrai, même si ε est très petit.

Ainsi, le terme u_n, bien qu'il n'atteigne peut-être jamais la valeur 1, est « *aussi proche que l'on veut de* 1 » *pour n assez grand.*

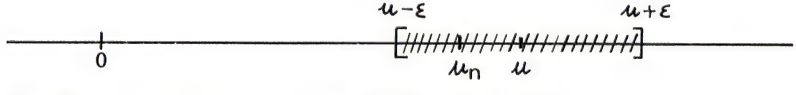
Plus précisément, pour tout $\varepsilon > 0$, il existe un entier n_0 tel que

$$n \geqslant n_0 \overset{\text{implique}}{\Longrightarrow} |u_n - 1| \leqslant \varepsilon$$

On traduit ce phénomène en disant que la suite u_n tend vers 1.

D'une manière générale, on dit qu'*une suite* u_n *tend vers un réel* u si pour tout $\varepsilon > 0$ il existe un entier n_0 tel que

$$n \geqslant n_0 \Longrightarrow |u_n - u| \leqslant \varepsilon \text{ ou } u - \varepsilon \leqslant u_n \leqslant u_n + \varepsilon$$
$$\text{ou } u_n \in [u - \varepsilon, u + \varepsilon]$$



On dit encore que u_n a u pour limite. On écrit :

$$u_n \longrightarrow u \quad \text{ou} \quad \lim_{n \to \infty} u_n = u$$

Une suite u_n qui a une limite u est dite *convergente.*

Une suite u_n qui n'a pas de limite est dite *divergente.*

La suite $u_n = (-1)^n$ est divergente, puisque, quel que soit l'entier n, si n est pair $u_n = +1$, si n est impair $u_n = -1$. Elle ne peut donc pas tendre vers $+1$, ni vers -1, ni vers aucun autre réel.

La suite $u_n = \frac{3}{2} n$ est aussi divergente. Mais le mode de divergence est différent du précédent. u_n « devient très grand » quand n augmente. Choisissons un réel $A > 0$. La propriété $u_n \geqslant A$ est vraie pour n assez grand, même si A a été choisi très grand.
On traduit ce phénomène en disant que u_n tend vers plus l'infini ; on écrit : $u_n \longrightarrow + \infty$

Plus précisément, on dit qu'*une suite* $u_n \longrightarrow + \infty$ si pour tout $A > 0$ il existe un entier n_0 tel que :

$$n \geqslant n_0 \longrightarrow u_n \geqslant A$$

De même, on dit qu'*une suite* $u_n \longrightarrow - \infty$ si u_n est inférieure à tout nombre négatif donné pour n assez grand, c'est-à-dire si pour tout $A \leqslant 0$ il existe un entier n_0 tel que

$$n \geqslant n_0 \longrightarrow u_n \leqslant A$$

Ainsi, la suite $u_n = 10 - n$ tend vers $- \infty$.

Détermination de la limite d'une suite

Il est bien évident qu'une suite *ne peut avoir qu'une limite finie ou infinie.*
La limite d'une suite se détermine le plus souvent en utilisant les théorèmes suivants :

1. Si $u_n \longrightarrow u$, *si* $v_n \longrightarrow v$, alors $u_n + v_n \longrightarrow u + v$ et $u_n \times v_n \longrightarrow u \times v$.
2. Si $u_n \longrightarrow u$ non nul, alors $\frac{1}{u_n} \longrightarrow \frac{1}{u}$.

Ces deux théorèmes ne sont valables que pour les limites finies. Des règles existent dans le cas des limites infinies, elles sont analogues à celles que nous énoncerons pour les limites de fonctions (page 121).

Exemples de calcul de limites

Considérons les suites :

Suite arithmétique

Une suite arithmétique se construit de la façon suivante : chaque élément u_n s'obtient en ajoutant toujours le même nombre b à l'élément précédant u_{n-1}

$$u_1 = 1, u_2 = 1 + 3 = 4, u_3 = 4 + 3 = 7, u_4 = 7 + 3 = 10,$$

etc. est une suite arithmétique.
La suite $u_1 = b, u_2 = b + b = 2b, u_3 = 2b + b = 3b, \ldots, u_n = n \times b$ est une suite arithmétique.

$\boxed{u_n = b.n}$

$bn \longrightarrow +\infty$ si $b > 0$
$bn \longrightarrow -\infty$ si $b < 0$

Suite géométrique

Une suite géométrique se construit de la façon suivante : chaque élément u_n s'obtient en multipliant l'élément précédent par le même nombre x.
Ainsi $u_1 = 2, u_2 = 2 \times 7 = 14, u_3 = 14 \times 7 = 98, u_4 = 98 \times 7$, etc. est une suite géométrique.
La suite $u_1 = x, u_2 = x \times x = x^2, u_3 = x \times x \times x = x^3, \ldots, u_n = x^n$ est une suite géométrique.

$\boxed{u_n = x^n}$ où x est un réel donné.

Si $x = 0$, u_n est égal à la suite stationnaire 0

$u_n \longrightarrow 0$

Si $x = 1$, pour tout entier n, $u_n = 1$

$u_n \longrightarrow 1$

Si $x = -1$, pour tout entier n, $u_n = (-1)^n$

u_n n'a pas de limite.

Distinguons alors deux cas : $|x| > 1$ et $|x| < 1$.

Si $|x| > 1$, $|x| = 1 + h$, avec $h > 0$.

Par suite :

$$|u_n| = |x|^n = \underbrace{(1+h)(1+h)\ldots(1+h)}_{n \text{ fois}} > n \times h \blacklozenge$$

◆ Voir Binôme (formule).

On en déduit que $|u_n| \longrightarrow +\infty$.

Si $|x| < 1$, on peut poser $x = \frac{1}{y}$, avec $|y| > 1$.

Par suite, $|y^n| \longrightarrow +\infty$ et $|u_n| = \frac{1}{|y^n|} \longrightarrow 0$.

Suite de nombres réels permettant de définir une suite d'intervalles fermés emboîtés

Considérons la suite définie, page suivante, par :

$$u_1 = 1$$

$$u_2 = 1 - \frac{1}{2}$$

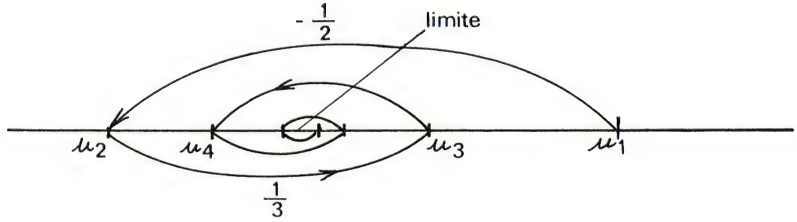
$$u_3 = 1 - \frac{1}{2} + \frac{1}{3}$$

$$u_4 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}$$

.....

$$u_n = 1 - \frac{1}{2} + \frac{1}{3} \ldots + \frac{(-1)^{n+1}}{n}$$

.....



Calculons effectivement u_1, u_2, u_3, u_4, u_5, u_6, ... Nous remarquons que tous les termes d'indices pairs sont inférieurs aux termes d'indices impairs et que :

$$u_2 \leqslant u_4 \leqslant u_6 \leqslant u_8 \leqslant \ldots \leqslant u_7 \leqslant u_5 \leqslant u_3 \leqslant u_1.$$

La suite d'intervalles

$$I_p = [u_{2p},\ u_{2p+1}]$$

est donc une suite d'intervalles fermés emboîtés dont la longueur tend vers 0 ; en effet,

$$u_{2p+1} - u_{2p} = \frac{(-1)^{2p+2}}{2p+1}$$

$$u_{2p+1} - u_{2p} \longrightarrow 0 \text{ quand } p \longrightarrow \infty.$$

L'intersection de cette suite d'intervalles est donc réduite à un point λ, et ce point λ est limite des deux suites

$$v_p = u_{2p} \qquad w_p = u_{2p+1}$$

Cela permet de démontrer que la suite $u_n \longrightarrow \lambda$.

Suite de nombres réels monotone

Une suite u_n est croissante ◆ si, pour tout entier n :

$$u_n \leqslant u_{n+1}$$

◆ Une suite stationnaire peut être considérée comme croissante et comme décroissante.

Une suite u_n est décroissante si, pour tout entier n :

$$u_n \geqslant u_{n+1}$$

Une suite u_n est monotone si elle est croissante ou décroissante.

Ainsi, la suite $u_n = \frac{1}{n}$ est décroissante ; la suite $u_n = 2n$ est croissante. Si $x > 1$, la suite $u_n = x^n$ est croissante ; si $0 < x < 1$, la suite $u_n = x^n$ est décroissante ; pour tout $x \geqslant 0$, la suite $u_n = x^n$ est monotone.
A toute suite u_n on peut associer l'ensemble des valeurs prises par cette suite. Nous noterons cet ensemble

$$\{ u_n \mid n \in \mathbb{N}^* \} \blacklozenge.$$

◆ Lire : ensemble des nombres u_n pour tout n appartenant à l'ensemble des nombres entiers positifs (0 exclu).

Par exemple, pour la suite u_n définie par $u_n = 2n$, l'ensemble $\{ u_n \mid n \in \mathbb{N}^* \}$ est l'ensemble des entiers pairs strictement positifs. Pour la suite u_n définie par $u_n = (-1)^n$, l'ensemble $\{ u_n \mid n \in \mathbb{N}^* \}$ est formé des seuls points $-1, +1$:

$$\{ u_n \mid n \in \mathbb{N}^* \} = \{ -1, +1 \}.$$

Une suite u_n est dite majorée (respectivement *minorée*, respectivement *bornée*) si l'ensemble $\{ u_n \mid n \in \mathbb{N}^* \}$ est majoré (respectivement minoré, respectivement borné).
Ainsi, la suite $u_n = 2n$ est minorée, elle n'est pas majorée. La suite $u_n = (-1)^n$ est majorée et minorée, donc également bornée.
On peut alors démontrer le théorème suivant relatif aux suites croissantes.

Si une suite u_n croissante est majorée, alors elle est convergente et

$$\lim u_n = \sup \{ u_n \mid n \in \mathbb{N}^* \} \blacklozenge$$

◆ Voir page 106 la définition de la borne supérieure d'un ensemble.

Si une suite u_n croissante n'est pas majorée, alors

$$\lim u_n = +\infty$$

Nous avons évidemment un théorème analogue pour les suites décroissantes :
Si une suite u_n décroissante est minorée, alors elle est convergente et

$$\lim u_n = \inf \{ u_n \mid n \in \mathbb{N}^* \}$$

Si une suite u_n décroissante n'est pas minorée, alors

$$\lim u_n = -\infty$$

Donnons un exemple d'utilisation de ces théorèmes sur les suites monotones. Considérons la suite u_n définie par

$$u_1 = \frac{1}{3} \qquad u_{n+1} = \frac{1}{2}(u_n^2 + 1)$$

Cette suite est croissante. En effet, pour tout entier n :

$$u_{n+1} - u_n = \frac{1}{2}(u_n^2 + 1) - u_n = \frac{1}{2}(u_n - 1)^2 \geqslant 0$$

Par ailleurs, cette suite est majorée.
Pour tout entier n, $u_n \leqslant 1$.
Cela résulte de ce que d'une part $0 \leqslant u_1 \leqslant 1$

d'autre part si $0 \leqslant u_n \leqslant 1$ alors $0 \leqslant u_{n+1} \leqslant 1$

Cela nous permet d'affirmer que la suite u_n est convergente.
On peut alors calculer sa limite u, car si $u_n \longrightarrow u$, on a aussi

$$u_{n+1} \longrightarrow u, \frac{1}{2}(u_n^2 + 1) \longrightarrow \frac{1}{2}(u^2 + 1).$$

Par suite, $u = \frac{1}{2}(u^2 + 1)$, $u_2 - 2u + 1 = 0$, $(u - 1)^2 = 0$, $u = 1$.

FONCTION D'UNE VARIABLE RÉELLE A VALEURS RÉELLES ◆

◆ Nous dirons, plus simplement, « fonction ».

Pour *définir une fonction f* d'une variable réelle à valeurs réelles, il faut se donner :
— une partie D de $\mathbb{R}$ (ensemble des nombres réels) appelée *domaine de définition de la fonction f* ;
— une règle qui à chaque élément x de D permet d'associer un nombre réel. Ce réel, noté le plus souvent $f(x)$, est appelé *image de x par la fonction f* ou encore :

valeur prise par la fonction f pour la valeur x de la variable.

Une telle fonction est notée :

$$f : D \longrightarrow \mathbb{R}$$
$$x \longmapsto f(x) \blacklozenge$$

◆ Lire : l'application f définie sur D a valeur dans l'ensemble des nombres réels $\mathbb{R}$; à tout x de D elle associe le nombre réel $f(x)$.

On peut envisager ainsi les fonctions :

1) $f : \mathbb{R} \longrightarrow \mathbb{R}$
$x \longmapsto x$

Fonction définie sur l'ensemble des nombres réels tout entier ($\mathbb{R} \longrightarrow \mathbb{R}$) qui à tout x fait correspondre x lui-même. On l'appelle *fonction identité* sur $\mathbb{R}$.

2) $f : \mathbb{R} - \{0\} \longrightarrow \mathbb{R}$
$x \longmapsto \frac{1}{x}$

Fonction définie sur l'ensemble des nombres réels sauf 0, qui à tout x (sauf $x = 0$) associe $\frac{1}{x}$. C'est un exemple particulier de *fonction homographique*.

3) $f : \mathbb{R} \longrightarrow \mathbb{R}$
$x \longmapsto x^2$

Fonction définie sur l'ensemble des réels tout entier, qui à tout x associe le carré de x. C'est la *fonction puissance 2.*

4) $f : \mathbb{R} \longrightarrow \mathbb{R}$
$x \longmapsto \mu$

Fonction définie sur l'ensemble des nombres réels tout entier, qui à tout x associe le même nombre réel μ. C'est la *fonction constante* égale à μ.

Ces fonctions sont souvent désignées de façon plus condensée par fonction :

$$f(x) = x \qquad f(x) = \frac{1}{x} \qquad f(x) = x^2 \qquad f(x) = \mu$$

ou encore par fonction :

$$y = x \qquad y = \frac{1}{x} \qquad y = x^2 \qquad y = \mu$$

L'inconvénient de ces notations est que le domaine de définition de la fonction n'apparaît pas. Les fonctions :

$$f : \mathbb{R} \longrightarrow \mathbb{R} \qquad f^* : [0, +\infty[\longrightarrow \mathbb{R} \blacklozenge$$
$$x \longrightarrow x^2 \qquad x \longrightarrow x^2$$

sont deux êtres mathématiques tout à fait différents ◆.

◆ La fonction est définie seulement pour les nombres réels positifs, sur l'intervalle $[0, +\infty[$.

◆ Ainsi que nous le verrons, la fonction f^* admet une fonction réciproque (voir plus loin, page 00), f n'en admet pas. La fonction f^* est appelée restriction de f à $[0, +\infty[$.

Bien entendu, une fonction est souvent définie comme « somme », « différence », « produit », « inverse » ou « quotient » d'autres fonctions.
En toute rigueur, si f_1 et f_2 sont définies sur un même domaine D de $\mathbb{R}$, on appelle :

somme de f_1 et f_2 la fonction, notée $f_1 + f_2$, définie par :

$$f_1 + f_2 : \mathrm{D} \longrightarrow \mathbb{R}$$
$$x \longrightarrow f_1(x) + f_2(x)$$

produit de f_1 et f_2 la fonction, notée $f_1 \times f_2$, définie par :

$$f_1 \times f_2 : \mathrm{D} \longrightarrow \mathbb{R}$$
$$x \longrightarrow f_1(x) \times f_2(x)$$

différence de f_1 et f_2 la fonction, notée $f_1 - f_2$, définie par :

$$f_1 - f_2 : \mathrm{D} \longrightarrow \mathbb{R}$$
$$x \longrightarrow f_1(x) - f_2(x)$$

De même, la *fonction inverse d'une fonction f* définie sur un domaine D de $\mathbb{R}$ est la fonction, notée $\frac{1}{f}$, définie par :

$$\frac{1}{f}(x) = \frac{1}{f(x)}$$

Pour que $\frac{1}{f(x)}$ existe, il faut que $f(x)$ soit défini et $f(x) \neq 0$ ◆.

◆ On ne peut pas définir l'inverse de 0 dans les nombres réels.

Le domaine de définition D′ de $\frac{1}{f(x)}$ sera donc celui de $f(x)$ auquel on enlève les points x tels que $f(x) = 0$

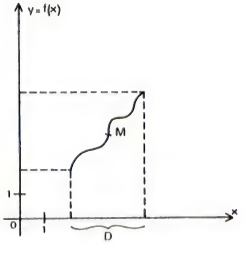
$$D' = \{ x \in D \mid f(x) \neq 0 \}$$

(lire : domaine de définition D′ = tous les x de D tels que $f(x)$ soit différent de 0).

REPRÉSENTATION D'UNE FONCTION

Donnons-nous deux axes de coordonnées rectangulaires Ox, Oy. Choisissons la même unité de longueur sur les deux axes et considérons une fonction :

$$f : D \longrightarrow \mathbb{R}.$$



Représentons sur l'axe Ox l'ensemble des points de D ; à tout point x de D nous associons le point M du plan d'abscisse x et d'ordonnée $y = f(x)$. L'ensemble des points M constitue la *représentation graphique de la fonction f*.
Il est à remarquer qu'à tout point x de D ne correspond qu'un seul point M. La représentation graphique d'une fonction est en général formée de un ou de plusieurs arcs de courbe que l'on peut matériellement tracer ◆.

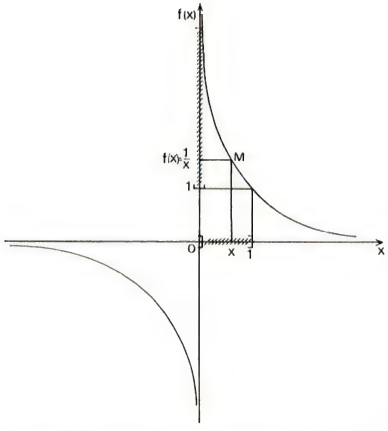
◆ Voir Fonctions.

IMAGE PAR UNE FONCTION D'UN ENSEMBLE

Considérons la fonction :

$$f : \mathbb{R} - \{ 0 \} \longrightarrow \mathbb{R}$$
$$x \longmapsto \frac{1}{x}.$$

C'est la fonction qui à tout nombre réel non nul x associe le nombre réel $\frac{1}{x}$, qui est son inverse. Ainsi, à 2 elle associe $\frac{1}{2}$; à 6, $\frac{1}{6}$; à 3, $\frac{1}{3}$ et à 0,2 elle associe $\frac{1}{0{,}2} = 5$. Plus généralement, pour cette fonction,



tout nombre réel strictement positif et inférieur ou égal à 1 a pour image son inverse, qui est un nombre réel supérieur ou égal à 1, et tout nombre réel supérieur ou égal à 1 peut être considéré comme l'inverse (donc l'image par f) d'un nombre réel strictement positif et inférieur ou égal à 1. La représentation graphique de la fonction $\frac{1}{x}$ permet de concrétiser ce que nous venons de dire.

Nous disons que l'intervalle ◆ $[1, +\infty[$ est l'image par f de l'intervalle $]0, 1]$ et nous écrivons : $f(]0, 1]) = [1, +\infty[$.

◆ Il est conseillé au lecteur de lire le mot Intervalle.

D'une manière générale, soit : $f : D \longrightarrow \mathbb{R}$.

Si E est une partie de D, on appelle *image de* E *par* f, et on note $f(E)$, l'ensemble formé par tous les réels y image d'au moins un élément x de E : $f(E) = \{ y \in \mathbb{R} \mid (\exists x \in E)\,(f(x) = y) \}$ ◆.

◆ Lire : l'image par f de E est l'ensemble de tous les nombres réels y, images par f d'un point x de E, quand x décrit tout E.

MAJORANT, MINORANT, BORNE

$f(E)$ est un ensemble de $\mathbb{R}$; il peut être majoré ◆, minoré, borné. Si $f(E)$ est majoré (respectivement minoré, respectivement borné), nous disons que la fonction f est *majorée sur* E (respectivement *minorée*, respectivement *bornée*). Mais, si l'ensemble $f(E)$ est majoré, il admet une infinité de majorants et nous savons, d'après le théorème de Bolzano-Weierstrass ◆, que parmi ses majorants il y en a un inférieur à tous les autres. Le plus petit majorant de $f(E)$ est appelé *borne supérieure de* f *sur* E et noté :

◆ Voir page 105.

◆ Voir page 106.

$$\sup_{x \in E} f(x) = \sup f(E).$$

De même, si l'ensemble $f(E)$ est minoré, il admet un plus grand minorant, appelé *borne inférieure de* f *sur* E et noté :

$$\inf_{x \in E} f(x) = \inf f(E).$$

Reprenons l'exemple de la fonction $f(x) = \frac{1}{x}$.

f est minorée sur $]0, +\infty[$ (voir graphique précédent), mais n'est pas majorée. Le plus grand minorant de f sur $]0, +\infty[$ est 0 :

$$\inf_{x > 0} f(x) = \inf_{x > 0} \left(\frac{1}{x}\right) = 0.$$

CONTINUITÉ ET LIMITE

CONTINUITÉ

Considérons une fonction f définie sur un ensemble D et x_0 un point de D. Supposons que x_0 ne soit pas isolé dans D : il y a des points de D proches de x_0, « très » proches de x_0, « aussi proches que l'on veut ».
Si x est proche de x_0, $f(x)$ est-il proche de $f(x_0)$? « Très » proche de $f(x_0)$? « Aussi proche que l'on veut » de $f(x_0)$?
Si l'on peut rendre $f(x)$ aussi proche que l'on veut de $f(x_0)$ en prenant x assez proche de x_0, on dit que $f(x)$ *est continue au point* x_0.

Quel que soit le nombre positif ε, il existe un nombre positif η tel que, si $|x - x_0| \leqslant \eta$, alors $|f(x) - f(x_0)| \leqslant \varepsilon$; ce que nous écrivons :

$$(\forall \varepsilon > 0)\ (\exists\, \eta > 0)\ (|x - x_0| \leqslant \eta \Longrightarrow |f(x) - f(x_0)| \leqslant \varepsilon).$$

Cela se dit encore : $f(x)$ *tend vers* $f(x_0)$ *lorsque* x *tend vers* x_0, ou : *la limite de* $f(x)$ *est* $f(x_0)$ lorsque x tend vers x_0. On écrit :

$$f(x) \longrightarrow f(x_0) \text{ lorsque } x \longrightarrow x_0$$

ou encore :

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Bien entendu, une fonction n'est pas toujours continue partout où elle est définie ◆.

◆ Une fonction est dite discontinue en un point a où elle est définie, si elle n'est pas continue en ce point.

LIMITE

Une notion plus générale que celle de continuité est celle de limite.
Une fonction f étant définie au *voisinage* ◆ d'un point a, on dit que $f(x)$ tend vers un nombre l lorsque x tend vers a, et on écrit :

$$f(x) \longrightarrow l \text{ lorsque } x \longrightarrow a$$

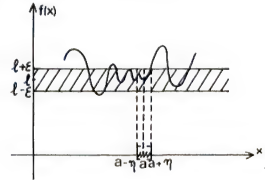
ou :

$$\lim_{x \to a} f(x) = l$$

si on peut rendre $f(x)$ aussi proche que l'on veut de l en prenant x assez proche de a.
Pour qu'il en soit ainsi, il faut et il suffit que, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que :

$$\begin{matrix} a - \eta \leqslant x \leqslant a + \eta \\ x \neq a \qquad x \in \mathrm{D} \end{matrix} \Longrightarrow |f(x) - l| \leqslant \varepsilon.$$

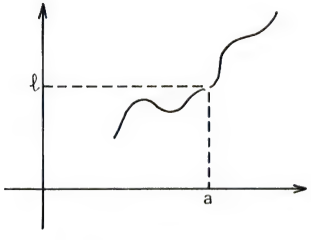
◆ f est définie au voisinage de a si, quel que soit le nombre positif η, il existe des points de D (ensemble de définition de f) vérifiant :
$|x - a| < \eta$
et :
$x \neq a$.
$f(x) = \dfrac{1}{x}$ est définie au voisinage de 0, pas en 0.
$f(x) = x^2$ est définie au voisinage de 0 et en 0.



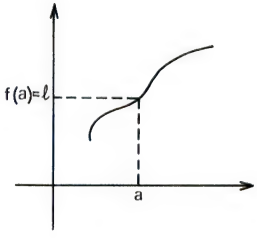
Ce qui s'écrit encore :

$$0 < |\, x - a \,| \leqslant \eta \atop x \in D \Longrightarrow |\, f(x) - l \,| \leqslant \varepsilon.$$

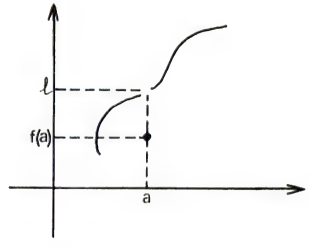
Une fonction peut ne pas avoir de limite quand x tend vers a (nous en verrons rapidement quelques exemples plus loin).
Elle peut avoir une limite quand x tend vers a. Alors, trois cas peuvent se produire :



Cas de discontinuité :
f n'est pas définie au point a.



Cas de continuité :
f est définie au point a ;
$l = \lim\limits_{x \to a} f(x) = f(a)$



Cas de discontinuité :
f est définie au point a ;
$l = \lim\limits_{x \to a} f(x) \neq f(a)$

PROPRIÉTÉS DES LIMITES ET DE LA CONTINUITÉ

● Si une fonction f a une limite lorsque x tend vers a, cette limite est unique.

● Si $f(x)$ tend vers l et $g(x)$ tend vers l' lorsque x tend vers a et si, pour tous les x assez proches de a, $f(x) \leqslant g(x)$, alors $l \leqslant l'$.

● Un cas particulier où une fonction n'a pas de limite quand x tend vers a est celui où l'on dit que $f(x)$ tend vers $+\infty$, un autre celui où l'on dit que $f(x)$ tend vers $-\infty$ ◆.

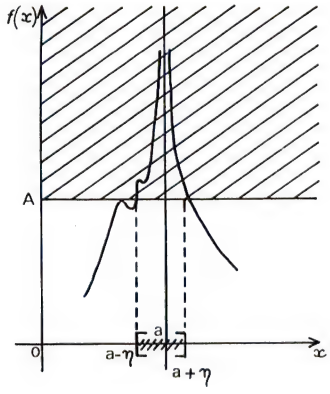
◆ Voir schéma n° 1, page suivante.

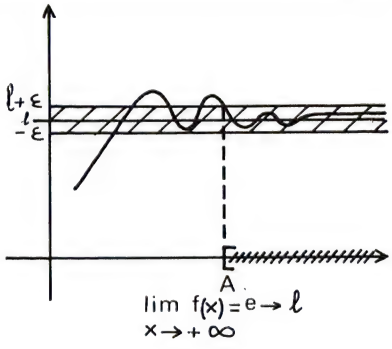
$f(x)$ tend vers $+\infty$ quand x tend vers a si, lorsque x est proche de a, $f(x)$ devient grand, « très » grand, « aussi grand que l'on veut ». Quel que soit le nombre positif A, il existera η positif tel que, si $|\, x - a \,| \leqslant \eta$, alors $f(x) \geqslant A$:

$$(\forall A > 0)\ (\exists\ \eta > 0)\ (x \in [a - \eta,\ a + \eta] \Longrightarrow f(x) \geqslant A).$$

De même, $f(x) \longrightarrow -\infty$ si :

$$(\forall B < 0)\ (\exists\ \eta > 0)\ (x \in [a - \eta,\ a + \eta] \Longrightarrow f(x) \leqslant B).$$





1er cas : $\lim_{x \longrightarrow a} f(x) \longrightarrow + \infty$ 2e cas : $\lim_{x \longrightarrow + \infty} f(x) \longrightarrow l$

● Cette approche de l'« infini » nous permet de définir également la limite d'une fonction quand $x \longrightarrow + \infty$ ($x \longrightarrow - \infty$ respectivement). Pour que x puisse tendre vers $+ \infty$, il est nécessaire que x puisse devenir « aussi grand qu'on veut », c'est-à-dire que l'ensemble de définition D de la fonction f ne soit pas borné supérieurement. l est la limite pour $x \longrightarrow + \infty$ de $f(x)$ si :

$$(\forall \varepsilon > 0)\ (\exists A > 0)\ (x \geqslant A \longrightarrow |f(x) - l| < \varepsilon)$$

(définition analogue pour $x \longrightarrow - \infty$).

● Nous avons aussi des propriétés de calcul des limites.
Si deux fonctions f et g ont des limites finies ou infinies, pour $x \longrightarrow a$ (ou $+ \infty$, ou $- \infty$), on peut, le plus souvent, en déduire la limite des fonctions :

$$f(x) + g(x) \qquad f(x) \times g(x) \qquad \frac{1}{f(x)}$$

Nous rassemblons ces résultats dans des tableaux :

si $f(x)$ tend vers	l	l	$+ \infty$	l	$- \infty$	$+ \infty$
et si $g(x)$ tend vers	l'	$+ \infty$	$+ \infty$	$- \infty$	$- \infty$	$- \infty$
alors $f(x) + g(x)$ tend vers	$l + l'$	$+ \infty$	$+ \infty$	$- \infty$	$- \infty$	?

On ne peut pas conclure si $f(x) \longrightarrow + \infty$ et si $g(x) \longrightarrow - \infty$; on dit que l'on a la *forme indéterminée* $+ \infty - \infty$ ◆.

◆ Ainsi pour $x \longrightarrow + \infty$, si : $f(x) = 3x$, $g(x) = -2x$, $f(x) + g(x) \longrightarrow + \infty$.
Si $f(x) = 3x$, $g(x) = -4x$, $f(x) + g(x) \longrightarrow - \infty$.
Si $f(x) = 3x$, $g(x) = -3x + 2$, $f(x) + g(x) \longrightarrow 2$.

si $f(x)$ tend vers	l	$l \neq 0$	∞	0
et si $g(x)$ tend vers	l'	∞	∞	∞
alors $f(x) \times g(x)$ tend vers	$l \times l'$	∞	∞	?

On ne peut pas conclure si $f(x) \longrightarrow 0$ et si $g(x) \longrightarrow \infty$ ◆ ; on dit que l'on a la *forme indéterminée* $0 \times \infty$ ◆.

◆ On dit que $f(x) \longrightarrow \infty$ si $|f(x)| \longrightarrow +\infty$.

◆ Ainsi pour $x \longrightarrow +\infty$ si :

$f(x) = \frac{1}{x}$, $g(x) = x$,

$f(x) \times g(x) \longrightarrow 1$.

$f(x) = \frac{1}{x}$, $g(x) = x^2$,

$f(x) \times g(x) \longrightarrow +\infty$.

$f(x) = \frac{1}{x}$, $g(x) = x$,

$f(x) \times g(x) \longrightarrow 0$.

si $f(x)$ tend vers	$l \neq 0$	0	∞
alors $\frac{1}{f(x)}$ tend vers	$\frac{1}{l}$	∞	0

● Il résulte immédiatement des propriétés énoncées pour les limites que, si deux fonctions f et g sont continues en un point, il en est de même pour les fonctions : $f+g$ $f \times g$ $\frac{1}{f}$ (sous réserve que $\frac{1}{f}$ soit définie au point $x_0, f(x_0) \neq 0$).

Ces propriétés nous permettent d'étudier la continuité des fonctions :

$$g : \mathbb{R} \longrightarrow \mathbb{R} \qquad f : \mathbb{R} - \{0\} \longrightarrow \mathbb{R}$$

$$x \longmapsto x^2 \qquad x \longmapsto \frac{1}{x}$$

Elles sont continues partout où elles sont définies ; en effet :

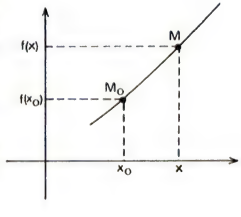
$$g(x) = x^2 = x \times x = \mathrm{I}(x) \times \mathrm{I}(x)$$

$$f(x) = \frac{1}{x} = \frac{1}{\mathrm{I}(x)}$$

où la fonction $\mathrm{I}(x) = x$ est bien évidemment continue en tout point x de $\mathbb{R}$.

● Nous venons de voir que les fonctions $g(x) = x^2$ et $f(x) = \frac{1}{x}$ étaient continues en tout point x_0 appartenant à l'ensemble de définition de ces fonctions. Quand une fonction est continue en tout points x_0 appartenant à un intervalle, on dit que *la fonction est continue sur cet intervalle.*

La représentation graphique d'une fonction continue sur un intervalle est formée d'un arc de courbe continue, car si $x \longrightarrow x_0$, alors $f(x) \longrightarrow f(x_0)$ et le point M de coordonnées $(x, f(x))$ tend vers le point M_0 de coordonnées $(x_0, f(x_0))$ quand $x \longrightarrow x_0$.



● Les fonctions continues sur un intervalle jouissent de propriétés particulièrement intéressantes.

1. — Si f est continue sur $[a, b]$, alors f est majorée et minorée ◆ sur $[a, b]$. De plus, elle atteint effectivement sa borne supérieure et sa borne inférieure.

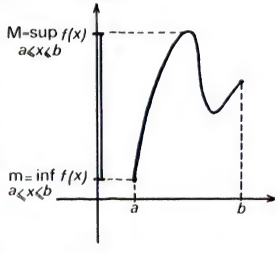
2. — Si f est continue sur $[a, b]$, alors f prend effectivement toute valeur y strictement comprise entre $f(a)$ et $f(b)$ en un point au moins de l'intervalle $[a, b]$.

De ces deux théorèmes, on déduit les corollaires suivants :

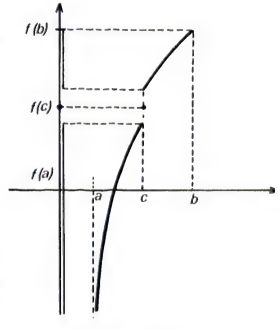
3. — Si f est continue sur $[a, b]$, $f([a, b]) = [m, M]$, où m et M désignent respectivement les bornes inférieure et supérieure de f sur $[a, b]$. On énonce encore : l'image par une fonction continue d'un intervalle fermé borné est un intervalle fermé borné.

4. — Enfin, l'image par une fonction continue d'un intervalle quelconque est un intervalle.

◆ Voir page 117.



f est continue sur $[a, b]$



f a une discontinuité en a et c

FONCTION CONTINUE STRICTEMENT MONOTONE FONCTION RÉCIPROQUE

Nous venons de voir que, si f est une fonction continue sur un intervalle I, l'image $f(I)$ de I est un intervalle. Nous allons avoir des résultats encore plus intéressants pour une fonction continue sur I et, de plus, strictement monotone sur I.

Une fonction est dite *monotone* sur un intervalle I (respectivement *strictement monotone*) si elle est croissante ou décroissante sur I (respectivement strictement croissante ou strictement décroissante).

Une fonction est croissante sur un intervalle I si, pour tout couple (x_1, x_2) d'éléments de I :

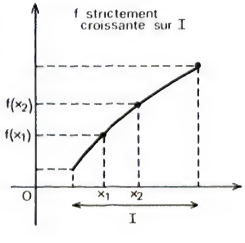
$$x_1 \leqslant x_2 \Longrightarrow f(x_1) \leqslant f(x_2)$$

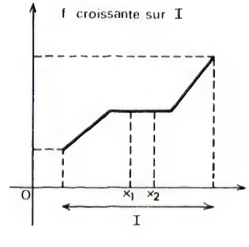
f est dite décroissante sur I si, de même, pour tout couple (x_1, x_2) d'éléments de I :

$$x_1 \leqslant x_2 \Longrightarrow f(x_1) \geqslant f(x_2)$$

On dit que f est strictement croissante (respectivement décroissante) si les inégalités sont strictes, c'est-à-dire si :

$$x_1 < x_2 \Longrightarrow f(x_1) < f(x_2) \text{ : strictement croissante}$$
$$x_1 < x_2 \Longrightarrow f(x_1) > f(x_2) \text{ : strictement décroissante}$$

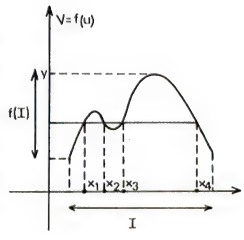




Revenons à f, une fonction définie continue sur un intervalle I de $\mathbb{R}$:

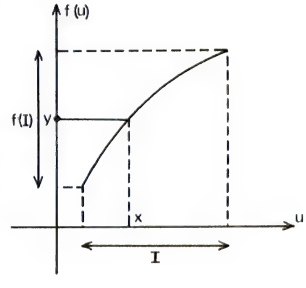
$$f : I \longrightarrow \mathbb{R}$$
$$x \longmapsto y = f(x)$$

Supposons f *continue sur* I. Nous avons vu que l'image $f(I)$ de I est un intervalle. Tout point y de $f(I)$ est donc l'image de un ou de plusieurs points de I. Sur la figure ci-contre, le point y est l'image de quatre points x_1, x_2, x_3, x_4 de I.



Mais, si f est de plus strictement monotone, alors tout point y de $f(I)$ est l'image d'un point x de I et d'un seul. Et alors, puisque tout élément y de $f(I)$ est l'image d'un élément de I et d'un seul, on peut envisager la fonction définie sur $f(I)$ qui à tout y de $f(I)$ associe l'élément x, unique, tel que $f(x) = y$.
Cette fonction est appelée *fonction réciproque* de la fonction f, notée f^{-1} :

$$f^{-1} : f(I) \longrightarrow \mathbb{R}$$
$$y \longmapsto x, \text{ unique, } x \in I, y = f(x).$$



Toute fonction f définie, continue, strictement monotone sur un intervalle I, *admet une fonction réciproque* f^{-1} *définie sur l'intervalle* $f(I)$ *par :*

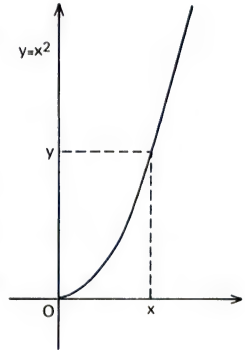
$$x = f^{-1}(y) \Longleftrightarrow \begin{cases} y = f(x) \\ x \in I. \end{cases}$$

Donnons un exemple. Nous considérons :

$$f : [0, +\infty[\longrightarrow \mathbb{R}$$
$$x \longmapsto y = x^2.$$

Comme nous l'avons vu, f est continue ; f est strictement croissante, sur $[0, +\infty[$:

$$f([0, +\infty[) = [0, +\infty[.$$



f admet donc une fonction réciproque f^{-1}, définie sur $[0, +\infty[$ par :

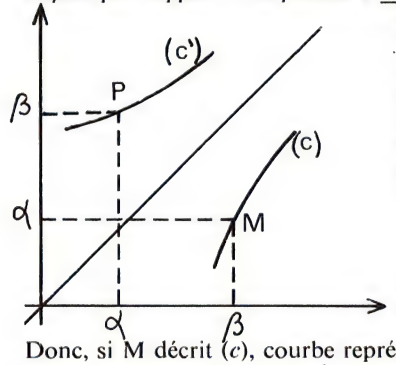
$$x = f^{-1}(y) \Longleftrightarrow \begin{cases} y = f(x) = x^2 \\ x \geqslant 0. \end{cases}$$

Connaissant y, on cherche x. Cette fonction est appelée *fonction puissance* $\frac{1}{2}$ ou encore *fonction racine carrée*. Elle se note $y^{\frac{1}{2}}$ ou $\sqrt{y}$.
Si nous reprenons la notation traditionnelle : y fonction de x, elle se note alors $x^{\frac{1}{2}}$ ou $\sqrt{x}$ et n'a de sens que pour $x \geqslant 0$:

$$y = x^{\frac{1}{2}} \Longleftrightarrow \begin{cases} x = y^2 \\ y \geqslant 0. \end{cases}$$

Essayons de dégager les propriétés de la fonction f^{-1}.

Les représentations graphiques (c) et (c') de f et f^{-1} relativement à un même système d'axes de coordonnées rectangulaires sont symétriques par rapport à la première bissectrice.



α étant un point quelconque de $f(I)$ et β un point quelconque de I, désignons par M le point de coordonnée (β, α) et par P le point de coordonnée (α, β). M et P sont symétriques par rapport à la première bissectrice. Or, si :

$$\alpha = f(\beta) \Longleftrightarrow \beta = f^{-1}(\alpha),$$

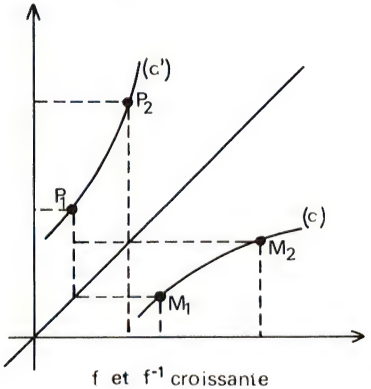
dans ce cas, M est un point de la courbe (*c*) et P de la courbe (*c'*).

Donc, si M décrit (*c*), courbe représentative de *f*, alors P décrit (*c'*), courbe représentative de f^{-1}, et réciproquement. C'est dire que (*c*) et (*c'*) sont symétriques par rapport à la première bissectrice.

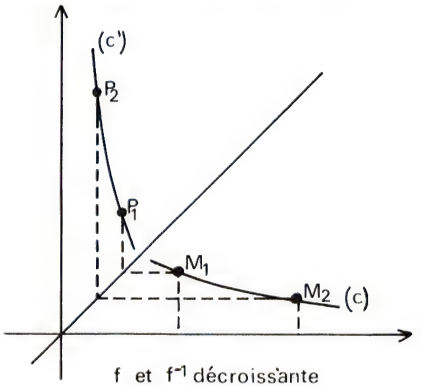
Cette symétrie des courbes (*c*) et (*c'*) nous permettra d'admettre que :

f^{-1} *est continue* (sa courbe représentative est, comme celle de *f*, formée d'un arc continu) ;

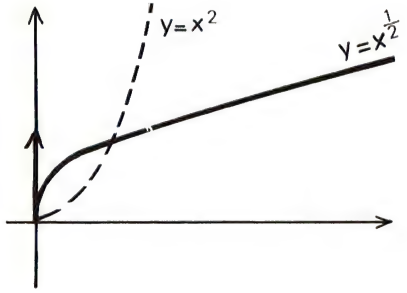
f^{-1} *est strictement monotone* et varie dans le même sens que *f*.



f et f^{-1} croissante



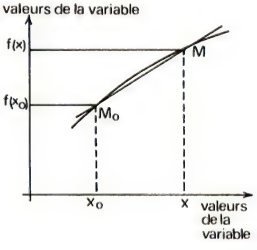
f et f^{-1} décroissante



Ainsi, la fonction puissance $\frac{1}{2}$ est continue, strictement croissante sur $[0, +\infty[$. Sa représentation graphique s'obtient par symétrie.

DÉRIVÉE D'UNE FONCTION ◆

◆ D'une variable réelle à valeurs réelles.



Soit f une fonction définie au voisinage d'un point x_0 et en x_0. On dit que *f est dérivable au point x_0* si le rapport $\dfrac{f(x) - f(x_0)}{x - x_0}$ a une limite finie lorsque x tend vers x_0. Cette limite, lorsqu'elle existe, est notée $f'(x_0)$ et appelée *dérivée de f au point x_0* ; on écrit :

$$f'(x_0) = \lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} \right)$$

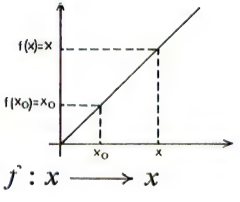
$\left(\text{lire : } f'(x_0) \text{ est égal à la limite, quand } x \text{ tend vers } x_0, \text{ de } \dfrac{f(x) - f(x_0)}{x - x_0}\right)$.

Fonction identité

Ainsi, la *fonction identité* définie sur $\mathbb{R}$ par $f(x) = x$ (fonction qui à tout x réel associe le nombre x) est dérivable au point x_0 quel que soit x_0.

En effet, pour tout $x \neq x_0$

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1$$

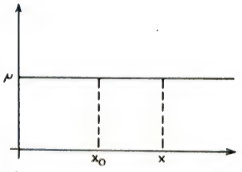


$f : x \longrightarrow x$

Par suite, $\lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = 1$. Donc f est dérivable au point x_0 et $f'(x_0) = 1$.

Fonction constante

De même, une fonction f constante sur $\mathbb{R}$ (à tout x de $\mathbb{R}$ f associe le même nombre réel $f(x) = \mu$) est dérivable en tout point x_0.



$f : x \longrightarrow \mu$

Pour tout $x \neq x_0$

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{\mu - \mu}{x - x_0} = 0$$

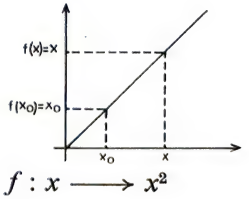
D'où $\lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = 0.$

Donc f est ici aussi dérivable au point x_0 et $f'(x_0) = 0$.

Fonction puissance 2

Enfin, si $f(x) = x^2$

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x - x_0)(x + x_0)}{x - x_0} = x + x_0$$



$f : x \longrightarrow x^2$

Et, lorsque $x \longrightarrow x_0$, alors $\dfrac{f(x) - f(x_0)}{x - x_0} \longrightarrow 2x_0$.

f est dérivable au point x_0 ; $f'(x_0) = 2x_0$.

Fonction puissance n

On montrerait de même, si n est un entier positif quelconque, que $f(x) = x^n = \underbrace{x \times x \times \ldots \times x}_{n \text{ fois}}$ est dérivable en tout point x_0 ; $f'(x_0) = nx_0^{n-1}$.

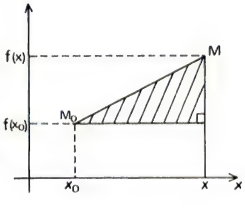
DÉRIVÉE INFINIE

Lorsque x tend vers x_0, le rapport $\dfrac{f(x) - f(x_0)}{x - x_0}$ n'a pas toujours une limite finie. Il peut arriver que le rapport

$$\frac{f(x) - f(x_0)}{x - x_0} \longrightarrow \pm \infty$$

Dans ce cas, si f est continue au point x_0, on dit que *f admet une dérivée infinie* au point x_0.

INTERPRÉTATION GÉOMÉTRIQUE



Nous pouvons donner une interprétation géométrique de la dérivée dans le cas d'une fonction f continue au point x_0. Soient (c) la représentation graphique de f relativement à un système d'axes Ox, Oy ; M_0 le point de coordonnées $\big(x_0, f(x_0)\big)$; M un point de coordonnées $\big(x, f(x)\big)$.

Faisons tendre x vers x_0, alors $f(x) \longrightarrow f(x_0)$ et par suite

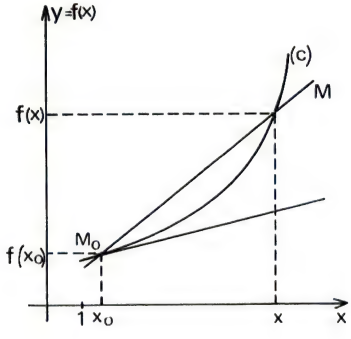
$$M_0M = \sqrt{(x - x_0)^2 + \big(f(x) - f(x_0)\big)^2} \longrightarrow 0.$$

Le point M tend vers le point M_0.

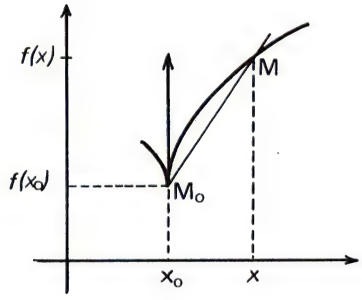
Dans ce cas, si la droite M_0M a une position limite :

Cette position limite est appelée *tangente à la courbe* (c) *en* M_0. Pour qu'il en soit ainsi, il faut et il suffit que f admette au point x_0 une dérivée finie ou infinie et la dérivée est alors égale à la *pente* de la tangente.

Cela résulte du fait que $\dfrac{f(x) - f(x_0)}{x - x_0}$ est égale à la pente de la droite M_0M.



dérivée finie



dérivée infinie (verticale)

Remarques : si une fonction est dérivable en un point x_0, alors elle est continue en ce point.

Si $\dfrac{f(x)-f(x_0)}{x-x_0} \longrightarrow f'(x_0)$ lorsque $x \longrightarrow x_0$ alors :

$$f(x)-f(x_0) = \frac{f(x)-f(x_0)}{x-x_0} \times (x-x_0) \longrightarrow f'(x_0) \times 0 = 0.$$

CALCUL D'UNE DÉRIVÉE

Nous savons déjà calculer en tous points x la dérivée des fonctions

$f(x) = \mu$	$f'(x) = 0$	$f(x) = x^2$	$f'(x) = 2x$
$f(x) = x$	$f'(x) = 1$	$f(x) = x^n$	$f'(x) = nx^{n-1}$

Nous pouvons en déduire les dérivées d'un grand nombre de fonctions :

$$f(x) = ax + b \qquad f(x) = \frac{1}{x} \qquad f(x) = \frac{1}{x^n}$$

en utilisant les deux théorèmes fondamentaux suivants.

Théorème 1 : dérivation d'une somme, d'un produit, d'un inverse.

Si deux fonctions f et g sont dérivables en un point x_0, il en est de même des fonctions

$$f+g \qquad f\times g \qquad \frac{1}{f}\blacklozenge$$

◆ Sous réserve que $\frac{1}{f}$ soit définie au point x_0, c'est-à-dire si $f(x_0)\neq 0$.

qui ont pour dérivées :

$$(f+g)'(x_0) = f'(x_0) + g'(x_0)$$

$$\boxed{(f+g)' = f' + g'}$$

dérivée d'une somme = somme des dérivées

$$(f\times g)'(x_0) = f(x_0)\,g'(x_0) + f'(x_0)\,g(x_0)$$

$$\boxed{(f\times g)' = fg' + gf'}$$

$$\left(\frac{1}{f}\right)'(x_0) = -\frac{f'(x_0)}{f^2(x_0)}$$

$$\boxed{\left(\frac{1}{f}\right)' = -\frac{f'}{f^2}}$$

Ces résultats sont des conséquences immédiates des théorèmes énoncés pour les limites pages 109 et 110 et des égalités :

- $$\frac{f(x)+g(x)-f(x_0)-g(x_0)}{x-x_0} = \frac{f(x)-f(x_0)}{x-x_0} + \frac{g(x)-g(x_0)}{x-x_0}$$
- $$\frac{f(x)\,g(x)-f(x_0)\,g(x_0)}{x-x_0} = f(x)\,\frac{g(x)-g(x_0)}{x-x_0} + g(x_0)\,\frac{f(x)-f(x_0)}{x-x_0}$$
- $$\frac{\dfrac{1}{f(x)}-\dfrac{1}{f(x_0)}}{x-x_0} = \frac{-1}{f(x)\,f(x_0)}\times\frac{f(x)-f(x_0)}{x-x_0}$$

Les deux derniers résultats permettent encore d'énoncer que, *si f et g sont dérivables en x_0, il en est de même de la fonction* $\frac{f}{g}\blacklozenge$.

◆ Sous réserve que $\frac{f}{g}$ soit définie au point x_0, c'est-à-dire si $g(x_0)\neq 0$.

Nous écrirons $\frac{f(x)}{g(x)}$ sous la forme d'un produit dont nous savons calculer la dérivée :

$$\frac{f(x)}{g(x)} = f(x)\times\frac{1}{g(x)}$$

$$\left(\frac{f}{g}\right)'(x_0) = -f(x_0)\times\frac{g'(x_0)}{g(x_0)^2} + f'(x_0)\times\frac{1}{g(x_0)}$$

$$\left(\frac{f}{g}\right)'(x_0) = \frac{g(x_0)\,f'(x_0)-f(x_0)\,g'(x_0)}{g(x_0)^2}$$

Utilisons ces résultats pour calculer les dérivées de fonctions citées précédemment

$$\boxed{f(x) = ax + b}$$

$$(ax + b)' = (ax)' + (b)' = a(x)' + (a)'x + (b)'$$
$$= a \times 1 + 0 \times x + 0 = a$$

$$\boxed{f(x) = \frac{1}{x}} \quad (x \neq 0)$$

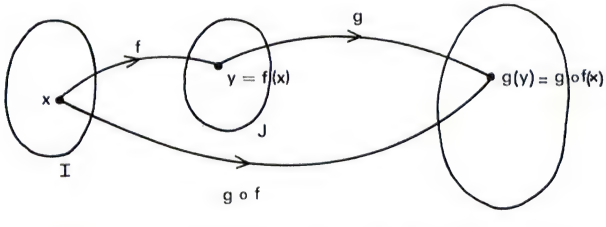
$$\left(\frac{1}{x}\right)' = -\frac{(x)'}{x^2} = -\frac{1}{x^2}$$

$$\boxed{f(x) = \frac{1}{x^n}} \quad (x \neq 0)$$

$$\left(\frac{1}{x^n}\right)' = -\frac{(x^n)'}{(x^n)^2} = \frac{-nx^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}}$$

Théorème 2 : dérivée d'une fonction composée

Considérons maintenant une fonction f définie sur un intervalle I et une fonction g définie sur un intervalle J. Si nous supposons $f(\mathrm{I}) \subset \mathrm{J}$, nous pouvons envisager la fonction notée $g \circ f$ (lire : g rond f) appelée *fonction composée de f par g* définie sur I par $g \circ f(x) = g[f(x)]$.



Nous avons alors la règle dite *règle de dérivation d'une fonction composée :*

Si f est dérivable au point x_0 et si g est dérivable au point $y_0 = f(x_0)$ correspondant, alors $g \circ f$ est dérivable au point x_0 de dérivée :

$$g \circ f'(x_0) = f'(x_0) \cdot g'(y_0)$$

Indiquons une démonstration valable lorsque au voisinage de x_0 $f(x) \neq f(x_0)$:

$$\frac{g \circ f(x) - g \circ f(x_0)}{x - x_0} = \frac{g\big(f(x)\big) - g\big(f(x_0)\big)}{x - x_0}$$

$$= \frac{g\big(f(x)\big) - g\big(f(x_0)\big)}{f(x) - f(x_0)} \times \frac{f(x) - f(x_0)}{x - x_0}$$

Donnons un exemple d'utilisation en envisageant la fonction F définie par $F(x) = (ax + b)^n$, où n est un entier positif. On peut considérer que $F = g \circ f$,

avec $f(x) = ax + b = y$ et $g(y) = y^n$.

f est dérivable en tout point x_0 : $f'(x_0) = a$.

g est dérivable au point $y_0 = f(x_0)$: $g'(y_0) = n{y_0}^{n-1}$.

On en déduit que F est dérivable en tout point x_0 et

$$F'(x_0) = an(ax_0 + b)^{n-1}$$

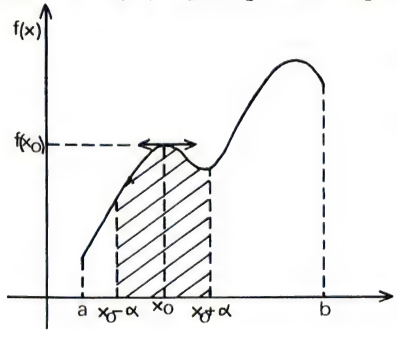
UTILISATION DE LA DÉRIVÉE

La notion de dérivée est fondamentale en analyse. On sait que l'étude de la dérivée d'une fonction permet de déterminer les *maximums et minimums relatifs* de cette fonction ; elle permet aussi, éventuellement, de déterminer les *intervalles de monotonie* de cette fonction.

Maximum et minimum

Qu'appelons-nous maximum ou minimum relatif d'une fonction? Une fonction f étant définie en x_0 et au voisinage de x_0, on dit que f admet un *maximum relatif* ◆ au point x_0 s'il existe un intervalle $[x_0 - \alpha, x_0 + \alpha]$ tel qu'en tout point de cet intervalle $f(x) \leqslant f(x_0)$.

◆ Si une fonction f définie sur un intervalle I a une borne supérieure sur I, effectivement atteinte en un point *intérieur* à I, f admet en ce point un maximum relatif.



Si une fonction f admet un maximum relatif en un point x_0 et si elle est dérivable en x_0, alors $f'(x_0) = 0$.

Nous considérons f admettant un maximum relatif en x_0. Il existe $\alpha > 0$ tel que :

$$x_0 - \alpha \leqslant x \leqslant x_0 + \alpha \Longrightarrow f(x) - f(x_0) \leqslant 0$$

alors $\dfrac{f(x) - f(x_0)}{x - x_0} \leqslant 0$ à droite de x_0 (pour $x - x_0 \geqslant 0$)

$$\frac{f(x) - f(x_0)}{x - x_0} \geqslant 0 \text{ à gauche de } x_0 \text{ (pour } x - x_0 \leqslant 0)$$

Dans ces conditions, la limite de $\dfrac{f(x) - f(x_0)}{x - x_0}$ pour $x \longrightarrow x_0$ ne peut être que nulle ; autrement dit:

$$f'(x_0) = 0$$

On définit de façon analogue le *minimum relatif* ◆. Si une fonction f admet en un point x_0 un minimum relatif et si elle est dérivable en x_0, alors $f'(x_0) = 0$.

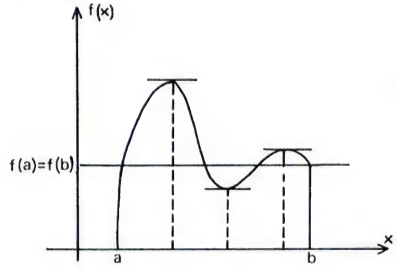
◆ Si une fonction définie sur un intervalle I a une borne inférieure effectivement atteinte en un point *intérieur* à I, f admet en ce point un minimum relatif.

Ainsi donc, si une fonction est dérivable en tout point d'un intervalle I, ses maximums et minimums relatifs devront être cherchés parmi les points de I en lesquels la dérivée est nulle.

Théorème de Rolle

Si f est une fonction définie, continue sur un intervalle $[a, b]$ dérivable sur $]a, b[$ ◆ avec $f(a) = f(b)$, alors il existe au moins un point de $]a, b[$ tel que $f'(c) = 0$.

◆ On appelle intervalle $[a, b]$ fermé l'ensemble des réels x tel que $a \leqslant x \leqslant b$, et intervalle ouvert $]a, b[$ l'ensemble des réels x tel que $a < x < b$ (strictement inférieur).



f est définie continue sur $[a, b]$. f est donc majorée et minorée sur $[a, b]$. Désignons par M et m ses bornes

$$\mathrm{M} = \sup_{x \in [a,\ b]} f(x) \qquad m = \inf_{x \in [a,\ b]} f(x)$$

Il est bien évident que :

soit $\mathrm{M} = m = f(a) = f(b)$

soit $\mathrm{M} > f(a)$ ou $m < f(a)$

Or si $\mathrm{M} = m = f(a) = f(b)$, alors f est constante sur $[a, b]$ et sa dérivée est nulle en tout point c de $]a, b[$.

Si $\mathrm{M} > f(a)$, f étant continue sur $[a, b]$, la valeur M est effectivement atteinte en un point c distinct de a et b et en ce point c en lequel f a un maximum relatif $f'(c) = 0$.

De même, si $m < f(a)$, la valeur m est effectivement atteinte en un point c distinct de a et b et en ce point c en lequel f a un minimum relatif $f'(c) = 0$.
Dans tous les cas, il existe au moins un point c de $]a, b[$ tel que $f'(c) = 0$.

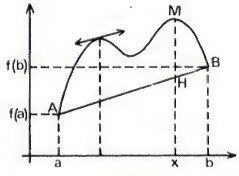
Théorème des accroissements finis (déduit du théorème de Rolle).

Si f est une fonction définie continue sur un intervalle $[a, b]$ dérivable sur $]a, b[$, il existe un point c de $]a, b[$ tel que

$$f(b) - f(a) = (b - a) f'(c)$$

La démonstration se fera en appliquant le théorème de Rolle à la fonction φ définie sur $[a, b]$ par

$$\varphi(x) = f(x) - f(a) - (x - a)\frac{f(b) - f(a)}{b - a}$$



Considérons les points A et B d'abscisse a, b de la représentation graphique (c) de f.
Si H et M sont les points de la droite AB et de la courbe (c) de même abscisse x : $\overline{HM} = \varphi(x)$.

Détermination des intervalles de monotonie ◆

Le théorème des accroissements finis permet d'établir les règles pratiques d'étude de la *monotonie des fonctions dérivables.*

◆ Voir la définition de monotonie, page 112.

Règle 1

Soit f une fonction dérivable sur un intervalle I. *Pour que f soit croissante sur* I, *il faut et il suffit qu'en tout point x de* I $f'(x) \geqslant 0$.

1) Supposons d'abord f croissante sur I. Quels que soient x et x_0 éléments de I

$$\frac{f(x) - f(x_0)}{x - x_0} \geqslant 0$$

Donc $\lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = f'(x_0) \geqslant 0.$

En tout point x_0 de I $f'(x_0) \geqslant 0$.

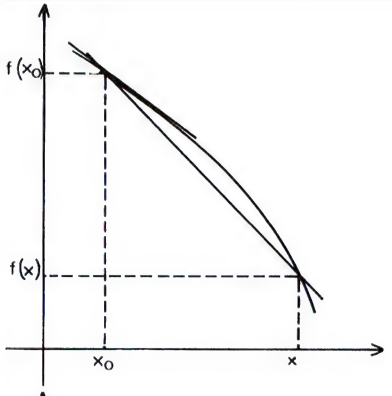
2) Réciproquement, supposons qu'en tout point x de I $f'(x) \geqslant 0$. Quels que soient x_1, x_2 éléments de I, f est dérivable sur $[x_1, x_2]$, donc continue sur $[x_1, x_2]$. Il existe un point c ◆ de $]x_1, x_2[$ tel que

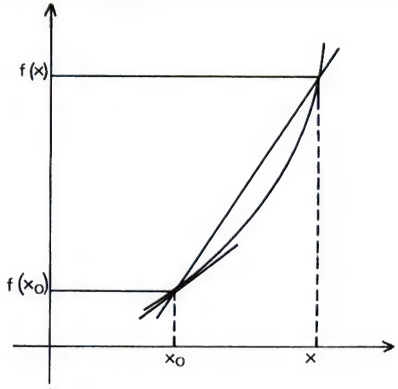
$$f(x_1) - f(x_2) = (x_1 - x_2) f'(c)$$

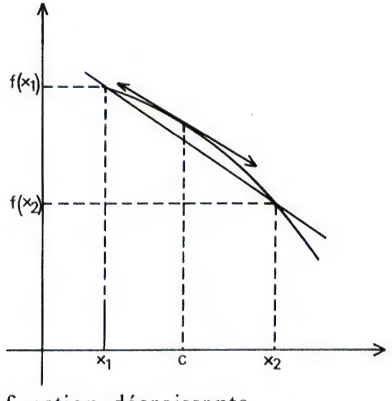
◆ D'après le théorème des accroissements finis ci-dessus.

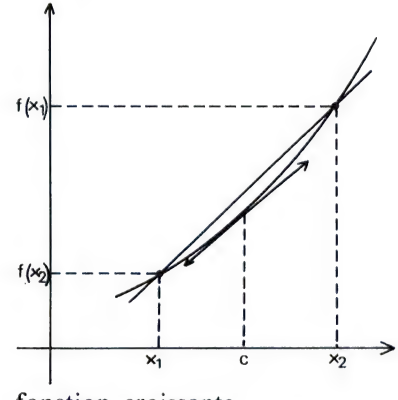
Mais $f'(c) \geqslant 0$. Donc $\dfrac{f(x_2) - f(x_1)}{x_2 - x_1} \geqslant 0.$

C'est dire que f est croissante sur I.
On démontrerait de même que f décroissante sur I équivaut à $f'(x) \leqslant 0$ en tout point x de I.





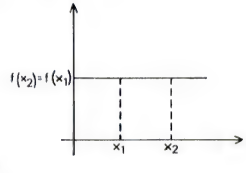




fonction décroissante — fonction croissante

$x_2 - x_1 > 0 \Longrightarrow f(x_2) - f(x_1) \leqslant 0$ $x_2 - x_1 > 0 \Longrightarrow f(x_2) - f(x_1) \geqslant 0$

On peut alors remarquer qu'une fonction constante sur un intervalle I est une fonction qui est à la fois croissante et décroissante sur I. On en déduit la règle 2.



Règle 2

Soit f une fonction dérivable sur un intervalle I. *Pour que f soit constante sur* I, *il faut et il suffit qu'en tout point x de* I $f'(x) = 0$.

Reprenons les fonctions citées en exemple au début du chapitre.

$$f(x) = \frac{1}{x}$$

f est dérivable en tout point $x \neq 0$, $f'(x) = -\frac{1}{x^2}$ ◆.

◆ Voir ci-dessus, page 129.

f est décroissante dans chacun des intervalles $]-\infty, 0[$ et $]0, +\infty[$.

$f(x) = x^2$

f est dérivable en tout point x $f'(x) = 2x$ ◆.

◆ Voir ci-dessus, page 127.

f est décroissante dans l'intervalle $]-\infty, 0]$

f est croissante dans l'intervalle $[0, +\infty[$.

ÉTUDE DE QUELQUES FONCTIONS

La méthode utilisée pour étudier complètement une fonction est exposée dans le dictionnaire sous le mot fonction.
L'étude détaillée des fonctions usuelles : fonction identité : $f(x) = x$, fonction constante : $f(x) = \mu$, fonction puissance : $f(x) = \cos x$, fonction exponentielle : $f(x) = e^x$, fonction logarithme : $f(x) = \text{Log } x$, se trouve dans le dictionnaire aux mots fonction, trigonométrie et logarithme.

Nous signalons au lecteur que le tableau des dérivées et des primitives se trouve page 313.

INTÉGRATION D'UNE FONCTION ◆

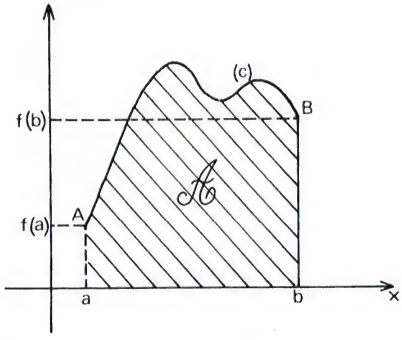
◆ Nous rappelons qu'il s'agit d'une variable réelle à valeurs réelles.

Dans cet article, nous nous proposons de donner quelques aperçus de la *théorie de l'intégration* après avoir défini ce que l'on appelle *fonction intégrable sur un intervalle* $[a, b]$ ◆.

◆ Lire d'abord ci-dessus, page 102, et Intervalle.

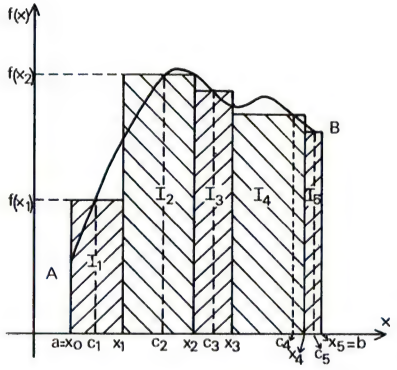
La notion d'intégrale est rattachée à la notion d'aire.

Considérons une fonction $f(x)$ définie, continue, positive sur un intervalle $[a, b]$.



Traçons sa courbe représentative (c). Nous ne pouvons en général calculer exactement l'aire $\mathcal{A}$, limitée par la courbe (c), les droites $x = a$, $x = b$ et l'axe des x.

Pour obtenir une valeur approchée de cette aire, nous pouvons par exemple diviser l'intervalle $[a, b]$ en cinq intervalles plus petits :



$$I_1 = [x_0, x_1]$$
$$I_2 = [x_1, x_2]$$
$$I_3 = [x_2, x_3]$$
$$I_4 = [x_3, x_4]$$
$$I_5 = [x_4, x_5]$$

Nous choisissons un point dans chacun de ces « petits intervalles ». Soient les points c_1, c_2, c_3, c_4, c_5.

La somme des rectangles :

$$(x_1 - x_0) f(c_1) + (x_2 - x_1) f(c_2) + (x_3 - x_2) f(c_3) + (x_4 - x_3) f(c_4) + (x_5 - x_4) f(c_5)$$

que nous notons encore

$$\sum_{i=1}^{i=n} (x_{i+1} - x_i) f(c_i) \blacklozenge$$

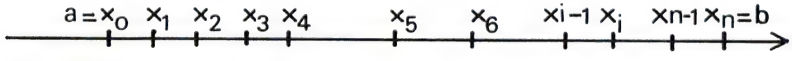
◆ Σ signifie somme de.

est égale à la somme des aires des rectangles hachurés sur notre figure et représente donc une valeur approchée de $\mathcal{A}$.
Évidemment, cette valeur approchée est très mauvaise.
Pour obtenir une valeur approchée meilleure, il faudrait diviser l'intervalle $[a, b]$ en un plus grand nombre de « petits intervalles » (et que la longueur du plus grand des « petits intervalles » soit assez petite).
Généralisons la démarche que nous venons d'effectuer.
On appelle *subdivision d'un intervalle* $[a, b]$, avec $a < b$, toute suite finie de points $x_0, x_1, x_2, \ldots, x_n$ tels que :

$$a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_{n-1} < x_n = b.$$

On appelle *module de la subdivision* σ formée par les points $x_0, x_1, x_2, \ldots, x_n$ et l'on note mod σ la longueur du « plus grand » des intervalles.
C'est la borne supérieure des n nombres $(x_1 - x_0), (x_2 - x_1), \ldots, (x_i - x_{i-1}), \ldots, (x_n - x_{n-1})$

$$\text{mod } \sigma = \sup_{1 \leqslant i \leqslant n} (x_i - x_{i-1})$$



Considérons maintenant :
— un intervalle $[a, b]$ quelconque, avec $a < b$;

— une subdivision σ de cet intervalle formée par les points $x_0 = a$, $x_1, x_2, \ldots, x_{i-1}, x_i, \ldots, x_n = b$;
— une fonction f définie, bornée sur $[a, b]$.
Dans chacun des intervalles $[x_{i-1}, x_i]$ de la subdivision σ, on peut choisir un point c_i et envisager la somme :

$$(x_1 - x_0) f(c_1) + (x_2 - x_1) f(c_2) + \ldots + (x_i - x_{i-1}) f(c_i) + \ldots + (x_n - x_{n-1}) f(c_n)$$

que l'on notera plus brièvement :

$$\sum_{i=1}^{i=n} (x_i - x_{i-1}) f(c_i)$$

Une telle somme s'appelle *somme de Riemann de la fonction f, relativement à la subdivision σ de l'intervalle $[a, b]$*.
Il est bien évident que, pour une même subdivision σ de l'intervalle $[a, b]$ et pour une même fonction f, il existe une infinité de sommes de Riemann possibles puisque, dans chaque intervalle $[x_{i-1}, x_i]$, il existe une infinité de façons de choisir le point c_i.
Les figures 1, 2, 3 montrent, pour une fonction f supposée continue et pour une même subdivision σ d'un intervalle $[a, b]$, différentes façons de choisir les poins c_i. Dans la figure 1, dans chaque intervalle $[x_{i-1}, x_i]$ le point c_i a été choisi au milieu de l'intervalle $[x_{i-1}, x_i]$:

$$c_i = \frac{x_{i-1} + x_i}{2}.$$

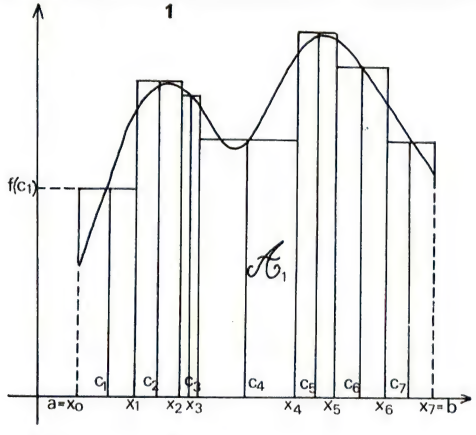
Dans la figure 2, dans chaque intervalle $[x_{i-1}, x_i]$, c_i a été choisi au point où la fonction atteint sa borne inférieure.
Dans la figure 3, dans chaque intervalle $[x_{i-1}, x_i]$, c_i a été choisi au point où la fonction atteint sa borne supérieure.

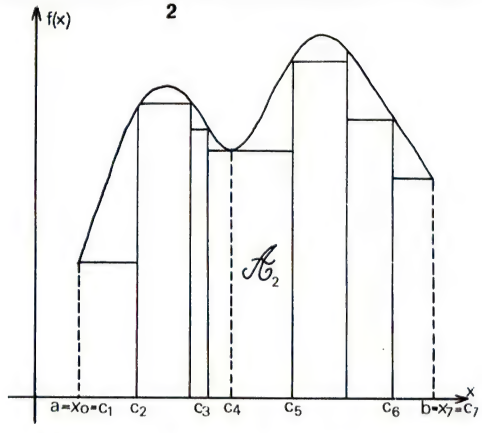
Nous remarquons sur ces figures que l'aire $\mathcal{A}$ est comprise entre les aires $\mathcal{A}_2$ et $\mathcal{A}_3$.

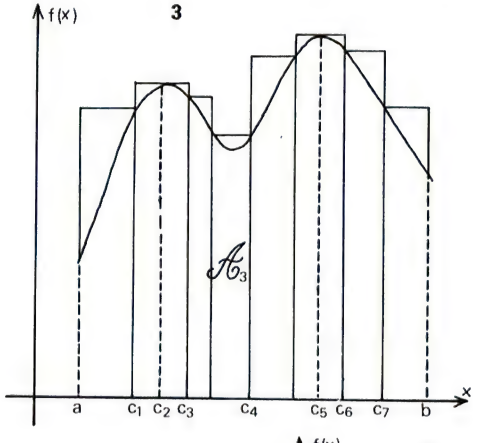
Si toutes les sommes de Riemann tendent vers une même limite I, lorsque le module des subdivisions « devient très petit », tend vers 0, le nombre I peut être considéré comme égal à $\mathcal{A}$.
D'où les définitions suivantes :
Soit f une fonction définie bornée sur $[a, b]$ avec $a < b$.
Nous notons :

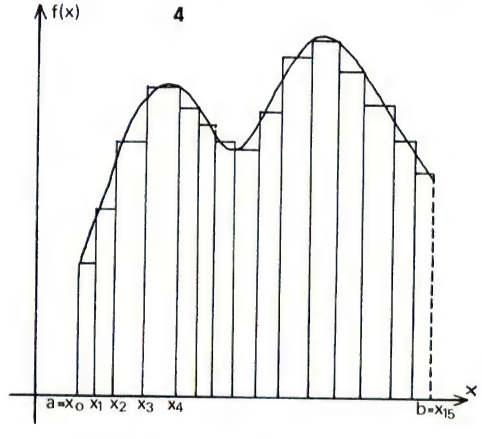
$$\Sigma\left(f, \sigma, (c_i)\right)$$

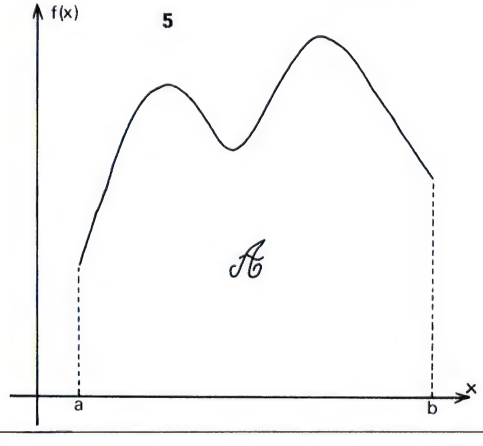
une somme de Riemann de f relativement à une subdivision σ de $[a, b]$ pour un choix particulier des points c_i. Nous disons que :
f est intégrable sur $[a, b]$ s'il existe un nombre réel I *limite commune à toutes les sommes de Riemann* $\Sigma\left(f, \sigma, (c_i)\right)$ lorsque le module de la subdivision σ tend vers 0.











$$\mathcal{A}_2 \leqslant \mathcal{A}_1 \leqslant \mathcal{A}_3$$
$$\mathcal{A}_2 \leqslant \mathcal{A} \leqslant \mathcal{A}_3$$

Plus précisément, f est intégrable sur $[a, b]$ ◆ si, pour tout nombre $\varepsilon > 0$, il existe $\eta > 0$ tel que :

$$\text{mod } \sigma \leqslant \eta \Longrightarrow | \Sigma (f, \sigma, (c_i)) - I | \leqslant \varepsilon$$

Comme on le vérifiera aisément, il ne peut exister deux tels nombres I. Lorsque I existe, I est appelée *intégrale de f sur* $[a, b]$ et notée :

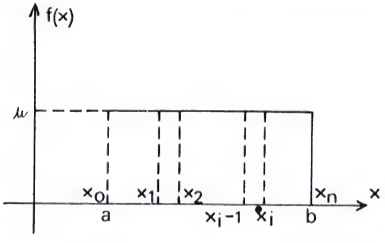
$$\boxed{\int_a^b f(x)\, dx}$$

(lire : intégrale de f entre a et b).
Cette notation peut paraître curieuse. Son intérêt apparaîtra lorsque nous étudierons la règle dite de « changement de variable ». Dans cette notation, la variable x ne joue aucun rôle et x peut être remplacée par une autre lettre :

$$I = \int_a^b f(x)\, dx = \int_a^b f(u)\, du = \int_a^b f(t)\, dt.$$

◆ Nous avons adopté le *point de vue de Riemann*. Riemann développa complètement au XIX^e siècle la théorie de l'intégration pour les *fonctions définies bornées sur un intervalle* $[a, b]$. En 1902, Lebesgue aborda d'une façon tout autre la théorie de l'intégration. La nouvelle méthode permettait d'« intégrer » un ensemble beaucoup plus vaste de fonctions. Toutefois, toute fonction bornée intégrable au sens de Riemann l'est au sens de Lebesgue et la valeur des deux intégrales est alors la même.

Exemples de fonctions intégrables



Considérons f constante sur $[a, b]$; $\boxed{f(x) = \mu.}$ Montrons que f est intégrable sur $[a, b]$ et calculons son intégrale.

Pour une subdivision σ particulière de $[a, b]$, quel que soit le choix des points c_i :

$$\begin{aligned}\Sigma (f, \sigma, (c_i)) &= \mu(x_i - x_0) + \mu(x_2 - x_1) + \ldots + \mu(x_i - x_{i-1}) + \ldots \\ &\quad + \mu(x_n - x_{n-1}) \\ &= \mu(x_1 - x_0 + x_2 - x_1 + \ldots + x_i - x_{i-1} + \ldots \\ &\quad + x_n - x_{n-1}) \\ &= \mu(x_n - x_0) = \mu(b - a)\end{aligned}$$

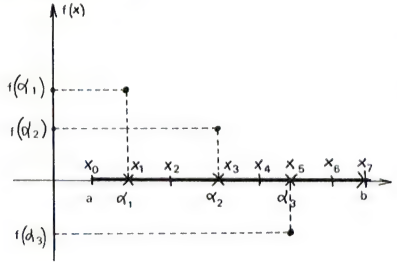
Toutes les sommes de Riemann ont la même valeur $\mu(b - a)$. $\mu(b - a)$ est la limite commune à toutes ces sommes de Riemann.

C'est dire que f est intégrable sur $[a, b]$ d'intégrale :

$$I = \int_a^b f(x)\, dx = \mu(b - a)$$

Résultat bien évident si on se réfère à l'interprétation géométrique d'une intégrale pour une fonction f continue (si μ est positif, $\mu(b - a)$ est la surface du rectangle hachuré).

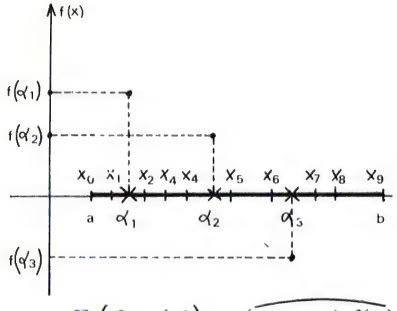
Donnons un deuxième exemple un peu moins trivial de fonctions intégrables sur un intervalle $[a, b]$ en considérant une *fonction nulle sur* $[a, b]$ *sauf en trois points* α_1, α_2, α_3 où elle prend les valeurs $f(\alpha_1)$, $f(\alpha_2)$, $f(\alpha_3)$.



Considérons d'abord une subdivision $\sigma = (x_0, x_1, x_2, \ldots, x_9)$ de $[a, b]$ dans laquelle les points α_1, α_2, α_3 sont, respectivement, intérieurs aux intervalles $[x_1, x_2]$, $[x_4, x_5]$, $[x_6, x_7]$. Dans une somme de Riemann associée à σ,

$$\Sigma\left(f, \sigma, (c_i)\right) = (x_1 - x_0) f(c_1) + \widehat{(x_2 - x_1) f(c_2)} + (x_3 - x_2) f(c_3)$$
$$+ (x_4 - x_3) f(c_4) + \widehat{(x_5 - x_4) f(c_5)} + (x_6 - x_5) f(c_6)$$
$$+ \widehat{(x_7 - x_6) f(c_7)} + (x_8 - x_7) f(c_8) + (x_9 - x_8) f(c_9)$$

Il y a au plus trois termes différents de 0, si c_2, c_5, c_7 sont précisément choisis en α_1, α_2, α_3.



Considérons maintenant une subdivision $\sigma = (x_0, x_1, x_2, \ldots, x_7)$ de $[a, b]$ dans laquelle les points α_1, α_2, α_3 coïncident respectivement avec x_1, x_3, x_5.
Dans une somme de Riemann associée à σ,

$$\Sigma\left(f, \sigma, (c_i)\right) = \widehat{(x_1 - x_0) f(c_1)} + \widehat{(x_2 - x_1) f(c_2)} + \widehat{(x_3 - x_2) f(c_3)}$$
$$+ \widehat{(x_4 - x_3) f(c_4)} + \widehat{(x_5 - x_4) f(c_5)} + \widehat{(x_6 - x_5) f(c_6)}$$
$$+ (x_7 - x_6) f(c_7)$$

Il pourra y avoir au plus $6 = 3 \times 2$ termes différents de 0 si c_1, c_2, c_3, c_4, c_5, c_6 sont respectivement choisis en α_1, α_1, α_2, α_2 α_3, α_3.
Dans tous les cas, quelle que soit la subdivision σ et quel que soit le choix des points c_i, dans une somme de Riemann $\Sigma\left(f, \sigma, (c_i)\right)$, il peut y avoir au plus six termes différents de 0.
En valeur absolue, chacun de ces six termes est inférieur ou égal à $A \times \text{mod } \sigma$, où A désigne le plus grand des trois nombres

$|f(\alpha_1)|$, $|f(\alpha_2)|$, $|f(\alpha_3)|$. On en déduit que, pour toute somme de Riemann $\Sigma(f, \sigma, (c_i))$, on a :

$$\left|\Sigma(f, \sigma, (c_i))\right| \leqslant 6\,A \bmod \sigma$$

C'est dire que 0 est limite commune à toutes les sommes de Riemann lorsque mod σ tend vers 0.

f est intégrable sur $[a, b]$ d'intégrale :

$$I = \int_a^b f(x)\,dx = 0$$

Aucune interprétation géométrique ne nous aurait permis de découvrir ce résultat, et la démonstration faite pour une fonction nulle en trois points se généralise immédiatement à une fonction nulle sur $[a, b]$ sauf en un nombre fini de points $\alpha_1, \alpha_2, \ldots, \alpha_n$.

On montre que *toute fonction continue sur un intervalle* $[a, b]$ *est intégrable sur* $[a, b]$.
On montre aussi qu'*une fonction monotone bornée sur un intervalle* $[a, b]$ *est intégrable sur* $[a, b]$.
Nous ne ferons pas ces démonstrations. Nous demanderons au lecteur de bien vouloir les admettre.

Généralisation du symbole $\int_a^b f(x)\,dx$

Revenons pour terminer ce paragraphe sur le symbole

$$\int_a^b f(x)\,dx$$

Pour $a < b$, le symbole a un sens si f est une fonction intégrable sur l'intervalle $[a, b]$; il désigne alors l'intégrale de f entre a et b. Pour $a \geqslant b$, nous n'avons jusqu'à présent pas donné de sens à ce symbole. Nous poserons *par définition*, f étant intégrable sur $[b, a]$:

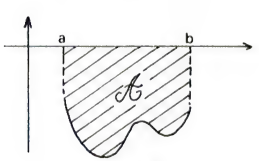
$$\text{si } a > b \qquad \int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

et, si $a = b$,

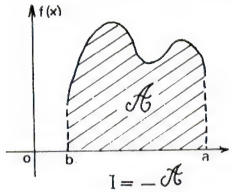
$$\int_a^a f(x)\,dx = 0$$

Il résulte de ces définitions que, quelle que soit la disposition relative des points a et b, f étant intégrable sur l'intervalle $[a, b]$ ou $[b, a]$:

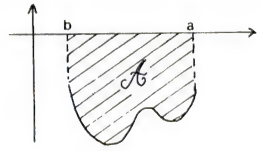
$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$



$I = -\mathcal{A}$



$I = -\mathcal{A}$



$I = +\mathcal{A}$

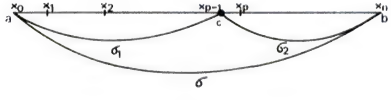
On peut alors, $\mathcal{A}$ désignant toujours l'aire limitée par l'axe des x, les droites $x = a$, $x = b$ et la courbe représentative de f, remarquer que si f garde un signe constant dans $[a, b]$, $I = \int_a^b f(x)\,dx = \pm\,\mathcal{A}$. Ceci, quel que soit les dispositions relatives de a et de b.

PROPRIÉTÉS DES FONCTIONS INTÉGRABLES

Propriétés relatives à l'intervalle d'intégration

Soient $[a, b]$ un intervalle, c un point tel que $a < c < b$, f une fonction bornée sur $[a, b]$.
Si f est intégrable sur $[a, c]$ et sur $[c, b]$, alors f est intégrable sur $[a, b]$:

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$



Nous considérons une fonction intégrable sur les intervalles $[a, c]$ et $[c, b]$ et nous nous proposons d'en déduire qu'elle est intégrable sur $[a, b]$.

Nous devons considérer une subdivision quelconque de $[a, b]$: $\sigma = (x_0, x_1, x_2, \ldots, x_n)$. Elle définit une subdivision σ_1 de $[a, c]$ formée des points $x_0, x_1, \ldots, x_{p-1}, c$ avec mod $\sigma_1 \leqslant$ mod σ, et une subdivision σ_2 de $[c, b]$ formée des points $c, x_{p+1}, \ldots, b$ et avec mod $\sigma_2 \leqslant$ mod σ.
L'étude de la somme de Riemann correspondante nous permet de démontrer le théorème énoncé.

Généralisons ce résultat.
Quel que soit l'ordre des points a, b, c ◆ :

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

◆ Sous réserve que chacun des symboles écrits ait un sens. Cela est d'ailleurs une condition suffisante.

Le théorème vient d'être démontré si $a < c < b$.

Il reste cinq cas à envisager :

$a \leqslant b \leqslant c \quad b \leqslant a \leqslant c \quad b \leqslant c \leqslant a \quad c \leqslant a \leqslant b \quad c \leqslant b \leqslant a$

Supposons, par exemple, $b \leqslant a \leqslant c$. D'après le résultat précédent :

$$\int_b^c f(x)\,dx = \int_b^a f(x)\,dx + \int_a^c f(x)\,dx$$

Alors :

$$\int_b^c f(x)\,dx - \int_a^c f(x)\,dx = \int_b^a f(x)\,dx$$

et, en multipliant les deux membres de l'égalité par -1 :

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx$$

Les autres cas se traitent de la même façon.

Propriétés de l'ensemble des fonctions intégrables

Nous allons maintenant étudier l'ensemble des fonctions intégrables sur un intervalle $[a, b]$.

La somme de deux fonctions intégrables sur un intervalle $[a, b]$ est une fonction intégrable sur $[a, b]$:

$$\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

Le théorème est vrai quel que soit l'ordre des points a et b. Indiquons la démonstration pour $a < b$.

Une somme de Riemann de $f + g$ sur $[a, b]$ relativement à une subdivision $a = x_0, x_1, \ldots, x_n = b$ peut s'écrire comme somme d'une somme de Riemann de f et d'une somme de Riemann de g.

$$\begin{aligned}(x_1 - x_0)\big(f(c_1) + g(c_1)\big) + (x_2 - x_1)\big(f(c_2) + g(c_2)\big) + \ldots \\ + (x_n - x_{n-1})\big(f(c_n) + g(c_n)\big) \\ = \underbrace{(x_1 - x_0)\,f(c_1) + \ldots + (x_n - x_{n-1})\,f(c_n)} \\ + \underbrace{(x_1 - x_0)\,g(c_1) + \ldots + (x_n - x_{n-1})\,g(c_n)}\end{aligned}$$

On en déduit le résultat annoncé par « passage à la limite ».

Si f est une fonction intégrable sur un intervalle $[a, b]$ et μ une constante réelle, alors μf est intégrable sur $[a, b]$ ◆ :

$$\int_a^b \mu f(x)\,dx = \mu \int_a^b f(x)\,dx$$

◆ L'ensemble des fonctions intégrables sur un intervalle $[a, b]$ constitue donc un espace vectoriel sur $\mathbb{R}$.

Nous laissons au lecteur le soin de faire cette démonstration qui pourra être conduite comme la précédente.

Propriétés de conservation des inégalités

Voyons maintenant la propriété des intégrales dite de « conservation des inégalités ». Cette propriété est conséquence de la propriété fondamentale suivante :

Si f est une fonction positive ou nulle intégrable sur un intervalle $[a, b]$ avec $a < b$, alors

$$\int_a^b f(x)\,\mathrm{d}x \geqslant 0$$

Toute somme de Riemann de f est positive ou nulle. Leur limite commune ne peut être que positive ou nulle.
La propriété de conservation des inégalités s'énonce alors :

Si f et g sont deux fonctions intégrables sur $[a, b]$, avec $a < b$, telles que pour tout x de $[a, b]$ $f(x) \leqslant g(x)$, alors:

$$\int_a^b f(x)\,\mathrm{d}x \leqslant \int_a^b g(x)\,\mathrm{d}x$$

On a : $\int_a^b \big(g(x) - f(x)\big)\,\mathrm{d}x \geqslant 0$

On en déduit le résultat, compte tenu du paragraphe précédent.

La propriété de conservation des inégalités a pour conséquence *la formule de la moyenne :*

Si f est une fonction continue sur un intervalle $[a, b]$, $a \neq b$, il existe au moins un point c de l'intervalle $[a, b]$ tel que :

$$\int_a^b f(x)\,\mathrm{d}x = f(c)\,(b - a)$$

Démontrons le résultat pour $a < b$.
f est continue sur $[a, b]$, donc f est majorée sur $[a, b]$ et minorée. Désignons par m sa borne inférieure et par M sa borne supérieure. Pour tout x de $[a, b]$:

$$m \leqslant f(x) \leqslant \mathrm{M}$$

On en déduit :

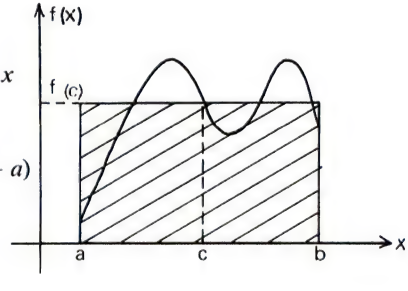
$$\int_a^b m\,\mathrm{d}x \leqslant \int_a^b f(x)\,\mathrm{d}x \leqslant \int_a^b \mathrm{M}\,\mathrm{d}x$$

soit :

$$m(b - a) \leqslant \int_a^b f(x)\,\mathrm{d}x \leqslant \mathrm{M}(b - a)$$

ou encore, en divisant par $b - a > 0$:

$$m \leqslant \frac{1}{b - a}\int_a^b f(x)\,\mathrm{d}x \leqslant \mathrm{M}$$



Le nombre réel $\frac{1}{b-a}\int_a^b f(x)\,dx$ est donc compris entre m et M.

Or les valeurs m, M, ainsi que toute valeur comprise entre m et M sont effectivement atteintes en un point au moins de l'intervalle $[a, b]$. Il existe un point c au moins de l'intervalle $[a, b]$ tel que :

$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx$$

PRIMITIVE ◆ *ET INTÉGRALE DES FONCTIONS CONTINUES SUR UN INTERVALLE*

◆ Voir également Primitive au dictionnaire.

Nous allons maintenant nous attacher à l'étude des seules fonctions continues et pour commencer établir la relation fondamentale existant entre intégrale et primitive d'une fonction continue sur un intervalle J quelconque.

> f étant une fonction définie sur un intervalle J, on appelle primitive de f sur J toute fonction F :
> — définie sur J ;
> — dérivable sur J, de dérivée $F'(x) = f(x)$ en tout point x de J.

Par exemple, $F(x) = \frac{1}{x}$ est une primitive de $f(x) = -\frac{1}{x^2}$ ◆.

◆ Voir la dérivée de $\frac{1}{x}$, page 129.

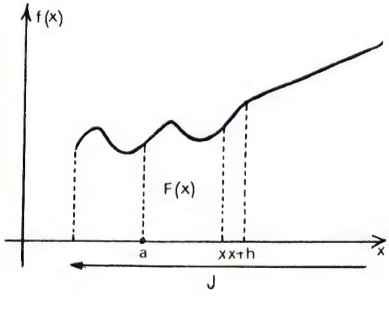
Théorème

Nous démontrons pour commencer le théorème suivant :

Si f est une fonction continue sur un intervalle J *quelconque, alors, quel que soit le point a de* J, *l'application :*

$$F : J \longrightarrow \mathbb{R}$$
$$x \longrightarrow \int_a^x f(t)\,dt$$

est dérivable sur J, *de dérivée* $F'(x) = f(x)$.



Choisissons un point a de J. Nous voulons montrer que l'application F définie sur J par :

$$F(x) = \int_a^x f(t)\,dt$$

est dérivable en tout point x de J. Pour cela, nous considérons :

Nous signalons au lecteur que le tableau des dérivées et des primitives se trouve page 313.

$$\frac{F(x+h)-F(x)}{h}=\frac{\left[\int_a^{x+h} f(t)\,dt-\int_a^x f(t)\,dt\right]}{h}=\frac{1}{h}\int_x^{x+h} f(t)\,dt$$

D'après la formule de la moyenne, quel que soit $h \neq 0$, il existe un point c compris entre x et $x+h$ tel que :

$$\frac{1}{h}\int_x^{x+h} f(t)\,dt = f(c)$$

Faisons alors tendre h vers 0 ; $x+h$ tend vers x, c compris entre x et $x+h$ tend aussi vers x et, f étant continue, $f(c) \longrightarrow f(x)$.

Ainsi, lorsque $h \longrightarrow 0$, $\dfrac{F(x+h)-F(x)}{h} \longrightarrow f(x)$.

C'est dire que F est dérivable au point x, de dérivée $F'(x) = f(x)$.
F est une primitive de f.
Ce théorème a deux corollaires.

Corollaire I *Toute fonction continue sur un intervalle* J *admet sur* J *une infinité de primitives. Pour que* Φ *soit une primitive de* f *sur* J, *il faut et il suffit que :*

$$\Phi(x) = \int_a^x f(t)\,dt + C$$

où C *est une constante arbitraire et* a *un point de* J.
Choisissons un point a de J d'après le théorème précédent.

$$F(x) = \int_a^x f(t)\,dt$$

est une primitive de f sur J. Il est alors évident qu'il en est de même de toute fonction $\Phi(x) = F(x) + C$, où C est une constante.
Réciproquement, supposons que Φ soit une primitive de f sur J. C'est dire que, pour tout x de J :

$$\Phi'(x) = f(x) = F'(x)$$

Soit encore que :

$$(\forall\, x \in J)\ \big(\Phi'(x) - F'(x) = 0\big).$$

Cela entraîne que :

$$\Phi(x) - F(x) = C \blacklozenge.$$

◆ Voir règle 2, page 133.

Corollaire II *Si* f *est continue sur* $[a, b]$ *et* Φ *une primitive particulière de* f *sur* $[a, b]$, *alors*

$$\int_a^b f(t)\,dt = \Phi(b) - \Phi(a)$$

Φ étant une primitive particulière de f sur $[a, b]$, désignons comme précédemment par F la primitive de F définie par $F(x) = \int_a^x f(t)\,dt$.

$$\Phi(x) = F(x) + C.$$

Mais $F(a) = 0$. Par suite, $C = \Phi(a)$ et

$$\Phi(x) - \Phi(a) = F(x)$$

et, en particulier, pour $x = b$:

$$\Phi(b) - \Phi(a) = F(b) = \int_a^b f(t)\,dt.$$

MÉTHODES PRATIQUES DE CALCUL D'UNE INTÉGRALE

Première méthode *Si on connaît une primitive particulière Φ de f sur $[a, b]$, on obtient immédiatement :*

$$\int_a^b f(x)\,dx = \Phi(b) - \Phi(a).$$

On note :

$$\Phi(b) - \Phi(a) = \Big[\Phi(x)\Big]_a^b.$$

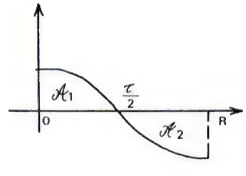
Premier exemple

sin x est une primitive de cos x sur $\mathbb{R}$. On en déduit :

$$\int_0^\tau \cos x\,dx = \Big[\sin x\Big]_0^\tau = \sin\tau - \sin 0 = 0.$$

Ce résultat était d'ailleurs prévisible géométriquement :

$$\int_0^\tau \cos x\,dx = \int_0^{\frac{\tau}{2}} \cos x\,dx + \int_{\frac{\tau}{2}}^\tau \cos x\,dx = \mathcal{A}_1 - \mathcal{A}_2$$



Second exemple

$\frac{x^3}{3}$ est une primitive particulière de x^2 sur $\mathbb{R}$.

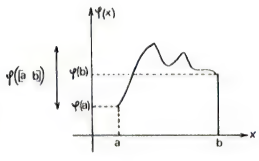
$$\int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right] = \frac{1}{3}.$$

Deuxième méthode *Supposons que nous ayons à calculer l'intégrale*

$$I = \int_a^b f\big(\varphi(x)\big)\,\varphi'(x)\,dx$$

φ étant une fonction continue de dérivée continue sur $[a, b]$, f étant continue sur l'intervalle $\varphi([a, b])$.

Si F est une primitive de f sur l'intervalle $\varphi([a, b])$, alors



$$F\big(\varphi(b)\big) - F\big(\varphi(a)\big) = \int_{\varphi(a)}^{\varphi(b)} f(t)\,dt.$$

Mais, par ailleurs, si F est une primitive de f sur l'intervalle $\varphi([a, b])$, d'après la règle de dérivation d'une fonction composée, $F\big(\varphi(x)\big)$ est une primitive de $f\big(\varphi(x)\big)\,\varphi'(x)$ et

$$F\big(\varphi(b)\big) - F\big(\varphi(a)\big) = \int_a^b f\big(\varphi(x)\big)\,\varphi'(x)\,dx.$$

D'où la règle dite de *changement de variable*. Dans les hypothèses énoncées, pour calculer

$$I = \int_a^b f[\varphi(x)]\,\varphi'(x)\,dx$$

on pose $t = \varphi(x)$ et $dt = \varphi'(x)\,dx$; alors

$$I = \int_{\varphi(a)}^{\varphi(b)} f(t)\,dt.$$

Premier exemple

Soit à calculer $\displaystyle\int_0^1 \frac{x}{(1+x^2)^2}\,dx$.

On pose $t = \varphi(x) = 1 + x^2$; alors $dt = \varphi'(x)\,dx = 2x\,dx$ et

$$\frac{1}{2}\int_0^1 \frac{2x}{(1+x^2)^2}\,dx = \frac{1}{2}\int_1^2 \frac{dt}{t^2} = \frac{1}{2}\left[-\frac{1}{t}\right]_1^2 = \frac{1}{4}.$$

Second exemple

Soit à calculer $\displaystyle\int_0^1 \sqrt{1-t^2}\,dt$.

Si on pose $t = \sin x$, alors $dt = \cos x\,dx$ et

$$\int_0^{\frac{\tau}{2}} \sqrt{1-\sin^2 x}\,\cos x\,dx = \int_0^1 \sqrt{1-t^2}\,dt.$$

Mais, pour $0 \leqslant x \leqslant \dfrac{\tau}{2}$, $\cos x \geqslant 0$, $\sqrt{1-\sin^2 x} = \sqrt{\cos^2 x} = \cos x$.

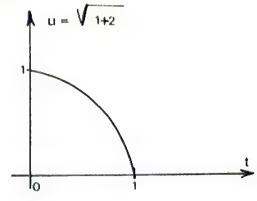
D'où : $\displaystyle\int_0^1 \sqrt{1-t^2}\,dt = \int_0^{\frac{\tau}{2}} \cos^2 x\,dx = \int_0^{\frac{\tau}{2}} \frac{\cos 2x + 1}{2}\,dx$

$$= \left[\frac{\sin 2x}{4} + \frac{x}{2}\right]_0^{\frac{\tau}{2}} = \frac{\tau}{4}.$$

Ce résultat était prévisible géométriquement. La courbe représentative de la fonction f définie dans [0, 1]

$$u = f(t) = \sqrt{1 - t^2}$$

est en effet un quart de circonférence.



$$u = \sqrt{1 - t^2} \iff \begin{cases} u^2 + t^2 = 1 \\ u \geqslant 0. \end{cases}$$

Troisième méthode *u et v étant des fonctions continues de dérivées continues sur l'intervalle* $[a, b]$

$$\int_a^b u(x)\, v'(x)\, dx = \Big[u(x)\, v(x)\Big]_a^b - \int_a^b v(x)\, u'(x)\, dx.$$

Cette règle, dite d'*intégration par partie*, est une conséquence immédiate de la règle de dérivation d'un produit.

$$[u(x)\, v(x)]' = u(x)\, v'(x) + v(x)\, u'(x)$$

$$\int_a^b [u(x)\, v(x)]'\, dx = \int_a^b u(x)\, v'(x)\, dx + \int_a^b v(x)\, u'(x)\, dx$$

$$= \Big[u(x)\, v(x)\Big]_a^b.$$

Exemple

Soit à calculer $\int_0^{\frac{\tau}{2}} x \cos x\, dx$.

Posons $u = x \quad u'(x) = 1 \quad v'(x) = \cos x \quad v(x) = \sin x$

$$\int_0^{\frac{\tau}{2}} x \cos x\, dx = \Big[x \sin x\Big]_0^{\frac{\tau}{2}} - \int_0^{\frac{\tau}{2}} \sin x\, dx$$

$$= \frac{\tau}{2} + \Big[\cos x\Big]_0^{\frac{\tau}{2}} = \frac{\tau}{2} - 1.$$

FONCTION DE DEUX OU TROIS VARIABLES

La notion de fonction de deux ou trois variables généralise immédiatement celle de fonction d'une variable. Ici, la variable n'est plus un réel, mais un couple de réels (ou un triplet). On note $\mathbb{R}^2$ (ou $\mathbb{R}^3$) l'ensemble des couples de réels (ou des triplets).
Une fonction f de deux (ou trois) variables, à valeurs réelles, est alors définie sur une partie D de $\mathbb{R}^2$ (ou de $\mathbb{R}^3$) ; elle associe à tout élément (x, y) (respectivement (x, y, z)) de cette partie un réel, noté $f(x, y)$ (respectivement $f(x, y, z)$).

Nous pouvons par exemple envisager les fonctions :

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} \qquad f : D \longrightarrow \mathbb{R}$$
$$(x, y) \longmapsto x + y \qquad (x, y) \longmapsto \sqrt{1 - x^2 - y^2}$$

La première est définie sur $\mathbb{R}^2$ tout entier, la deuxième est définie uniquement par les couples (x, y) tels que $x^2 + y^2 \leqslant 1$. Nous notons cet ensemble D :

$$D = \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1 \}.$$

Nous pouvons aussi envisager par exemple les fonctions :

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R} \qquad f : D \longrightarrow \mathbb{R}$$
$$(x, y, z) \longmapsto x + 2y + 3z \quad (x, y, z) \longmapsto \frac{6yz}{x - 2y}.$$

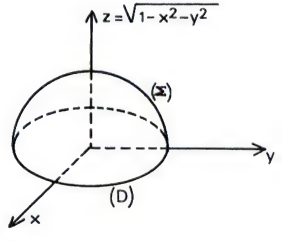
La première est définie sur $\mathbb{R}^3$ tout entier, la deuxième sur :

$$D = \{ (x, y, z) \in \mathbb{R}^3 \mid x - 2y \neq 0 \}.$$

REPRÉSENTATION

Comme nous avons représenté géométriquement une fonction d'une variable réelle, nous pouvons aussi représenter une fonction de deux variables définie sur un domaine D de $\mathbb{R}^2$.
Nous considérons un système de trois axes de coordonnées rectangulaires deux à deux. Nous choisissons la même unité de longueur sur les trois axes. L'ensemble des points M de coordonnées (x, y, z) avec $(x, y) \in D$ et $z = f(x, y)$ constitue la représentation graphique de f ◆.

◆ Elle est en général formée d'un morceau de surface Σ.



Reprenons l'exemple :

$$(x, y) \longmapsto \sqrt{1 - x^2 - y^2}$$

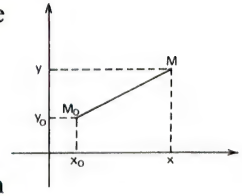
La représentation graphique de f est formée par la surface d'une demi-sphère :

$$z = \sqrt{1 - x^2 - y^2}$$
$$\Longleftrightarrow \begin{cases} z \geqslant 0 \\ x^2 + y^2 + z^2 = 1. \end{cases}$$

LIMITE, CONTINUITÉ

Si nous voulons généraliser la notion de limite, puis de continuité, il faut arriver à traduire la notion de proximité.
Dire qu'un point M de $\mathbb{R}^2$ tend vers un point M_0 de $\mathbb{R}^2$ signifie que la

distance entre M et M_0 tend vers 0. Soient (x, y) les coordonnées de M, (x_0, y_0) celles de M_0 ; nous définissons ainsi la distance entre M et M_0 : $d(M, M_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$.



De même, si M et M_0 sont deux points de $\mathbb{R}^3$, on pose :

$$d(M, M_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}.$$

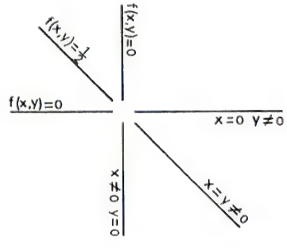
Soient f une fonction définie sur un domaine D de $\mathbb{R}^2$ (ou $\mathbb{R}^3$), M un point de D et M_0 un point de D ou de la frontière de D. On dit que f a une limite l lorsque M tend vers M_0 si $f(M)$ est aussi proche que l'on veut de l pour M assez proche de M_0, soit :

$$\forall \varepsilon > 0, \exists \eta > 0 \mid d(M, M_0) \leqslant \eta \Longrightarrow |f(M) - l| \leqslant \varepsilon.$$

Considérons maintenant f définie sur $\mathbb{R}^2$, sauf au point (0, 0), par :

$$f(x, y) = \frac{xy}{x^2 + y^2}.$$

(0, 0) n'appartient pas à D, domaine de définition de f, mais à la frontière de D. On peut chercher si $f(x, y)$ a une limite lorsque (x, y) tend vers (0, 0).
Pour tout couple (x, y) tel que $x = 0$, $y \neq 0$, $f(x, y) = 0$, de même que pour tout couple (x, y) tel que $x \neq 0$, $y = 0$.
Mais, pour tout couple (x, y) tel que $x = y$ avec $(x, y) \neq (0, 0)$:



$$f(x, y) = \frac{x \times x}{x^2 + x^2} = \frac{x^2}{2x^2} = \frac{1}{2}.$$

Il y a des couples « aussi proches que l'on veut » du couple (0, 0) tel que $f(x, y) = 0$. Il y a des couples « aussi proches que l'on veut » du couple (0, 0) tel que $f(x, y) = \frac{1}{2}$.

$f(x, y)$ n'a pas de limite lorsque $(x, y) \longrightarrow (0, 0)$.

Les propriétés des limites sont analogues à celles que nous avons vues pour les fonctions réelles à valeurs réelles.
On dit qu'une fonction f définie sur un domaine D de $\mathbb{R}^2$ (ou $\mathbb{R}^3$) est continue en un point A de D si $f(M)$ tend vers $f(A)$ quand M tend vers A, ce qui s'écrit encore :

$$\forall \varepsilon > 0, \exists \eta > 0 \mid d(M, A) \leqslant \eta \Longrightarrow |f(M) - f(A)| \leqslant \varepsilon.$$

La fonction que nous venons d'étudier, définie par :

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{et} \quad f(0, 0) = 0,$$

n'est pas continue au point (0, 0), car il existe des points M aussi proches que l'on veut de (0, 0) pour lesquels $f(M) = \frac{1}{2}$.

DÉRIVÉE PARTIELLE

DÉRIVÉE PREMIÈRE

Considérons une fonction f de deux variables réelles définie sur un domaine D de $\mathbb{R}^2$. Soit $M_0 = (x_0, y_0)$ un point appartenant à l'intérieur de D. Nous pouvons envisager la fonction F où seule la variable x varie, définie par :

$$F(x) = f(x, y_0)$$

parce que y_0 est fixé.
On dit que *f admet une dérivée partielle par rapport à x au point* (x_0, y_0) si F est dérivable au point x_0 ; elle se note $f'_x(x_0, y_0)$ et, par définition :

$$f'_x(x_0, y_0) = F'(x_0).$$

Nous pouvons de même envisager la fonction G de la seule variable y définie par $G(y) = f(x_0, y)$, où x_0 est fixé.
On dit que *f admet une dérivée partielle par rapport à y au point* (x_0, y_0) si G est dérivable au point y_0 ; elle se note $f'_y(x_0, y_0)$ et, par définition :

$$f'_y(x_0, y_0) = G'(y_0).$$

(La définition est analogue si f est une fonction de trois variables : alors, $f'_x(x_0, y_0, z_0) = F'(x_0)$ avec $F(x) = f(x, y_0, z_0)$, etc.)

DÉRIVÉE SECONDE

Lorsqu'une fonction admet une dérivée partielle en tout point d'un domaine D de $\mathbb{R}^2$ (ou $\mathbb{R}^3$), cette dérivée partielle définit une nouvelle fonction de deux (ou trois) variables qui peut éventuellement, à son tour, admettre une dérivée partielle. Par exemple, supposons que $f'_x(x, y)$ existe. Cette fonction peut être dérivable par rapport à x ; alors, cette nouvelle dérivée se note :

$$f''_{x^2} = (f'_x)'_x.$$

Elle peut être dérivable par rapport à y ; dans ce cas, cette dérivée se note :

$$f''_{xy} = (f'_x)'_y.$$

Une dérivée partielle de dérivée partielle est appelée dérivée seconde.
Il peut exister d'autres dérivées partielles secondes :

$$f''_{yx} \qquad f''_{y^2}$$

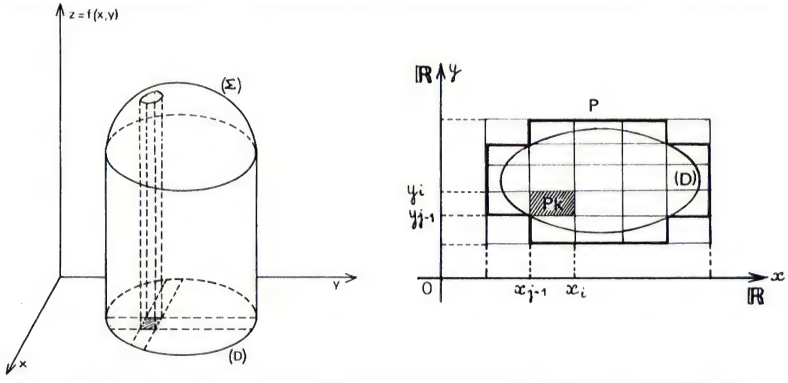
(Les définitions sont identiques pour trois variables.)

Ce qui est particulièrement remarquable, c'est que, en général, la valeur des dérivées secondes est indépendante de l'ordre dans lequel ont été effectuées les deux dérivations successives.
Ainsi, si f_{xy} et f_{yx} existent et sont continues au point (x_0, y_0) et au voisinage de (x_0, y_0) :

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

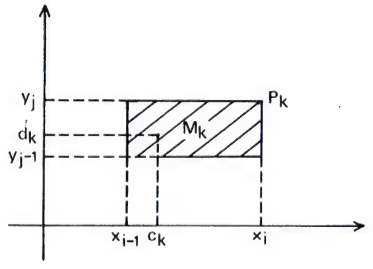
INTÉGRALE DOUBLE

La notion d'intégrale simple avait pu être introduite en s'appuyant sur la notion de surface. De même, nous allons introduire la notion d'intégrale double en nous appuyant sur la notion de volume. Nous avons vu qu'une fonction de deux variables réelles à valeurs réelles admet une représentation graphique dans $\mathbb{R}^3$.



Considérons une fonction f définie sur une partie D de $\mathbb{R}^2$.
Nous supposons D borné ; c'est dire que l'on peut inclure D dans un rectangle de côtés parallèles aux axes. Soit Σ la surface représentant la fonction $z = f(x, y)$. Supposons encore f positive sur D et proposons-nous de calculer le volume V du cylindre, de base D, limité par la surface Σ.
D est inclus dans un rectangle. Ce rectangle peut être fractionné en rectangles plus petits en le coupant par des parallèles aux axes. Considérons l'ensemble des petits rectangles: $P_1, P_2, \ldots, P_k, \ldots, P_p$, recouvrant D. Ils constituent un pavage $\mathcal{P}$ du domaine D. Calculons la surface de ces « petits rectangles ». Désignons par $\mu(P_k)$ la surface du rectangle P_k dont les sommets ont respectivement pour coordonnée x_{i-1}, x_i, y_{j-1}, y_j :

$$\mu(P_k) = (y_j - y_{j-1})(x_i - x_{i-1}).$$



Si on choisit alors un point :

$$M_k = (c_k, d_k),$$

dans chacun des rectangles P_k :

$$f(M_k) \times \mu(P_k)$$

représente « à peu près » le volume d'un « petit cylindre » de base P_k limité par Σ.

La somme de tous les volumes associés à tous les « petits » rectangles :

$$P_1, P_2, \ldots, P_k, \ldots, P_p$$

du pavage $\mathcal{P}$ constitue une valeur approchée du volume V.

Cette valeur approchée sera d'autant meilleure que les rectangles P_k seront plus petits. C'est une *somme de Riemann* de f relativement au pavage $\mathcal{P}$. Elle se note :

$$\Sigma\left(f, \mathcal{P}, (M_k)\right) = f(M_1)\,\mu(P_1) + f(M_2)\,\mu(P_2) + \ldots + f(M_k)\,\mu(P_k) + \ldots + f(M_p)\,\mu(P_p).$$

Plus généralement, f étant bornée sur D, mais pas nécessairement d'un signe constant, s'il existe un nombre réel I « limite commune, à toutes les sommes de Riemann » lorsque le module du pavage $\mathcal{P}$ ◆ tend vers 0, on dit que f est *intégrable sur D*. I est appelée *intégrale double de f sur D*, notée :

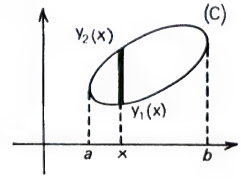
$$\iint_D f(x, y)\,dx\,dy.$$

◆ On appelle module du pavage $\mathcal{P}$ la surface du plus grand des rectangles $P_1, P_2, \ldots, P_k, \ldots, P_p$:

$$\text{mod } \mathcal{P} = \sup_{1 \leqslant k \leqslant p} P_k.$$

Une intégrale double se calcule, en général, par deux intégrations simples. Si le domaine D est limité par une courbe (C) qu'une parallèle à l'axe des y coupe en deux points d'ordonnée $y_1(x)$ et $y_2(x)$:

$$\iint_D f(x, y)\,dx\,dy = \int_a^b \left[\int_{y_1(x)}^{y_2(x)} f(x, y)\,dy\right] dx.$$

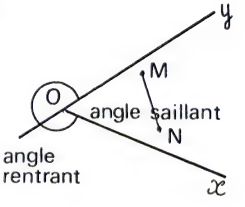


L'intégrale entre crochets se calcule comme si x était une constante.

Christiane de Bary.

ANGLE pages 14, 77, 94 et 95.

Angle géométrique (ou secteur angulaire) : deux demi-droites Ox, Oy, de même origine O, déterminent dans le plan deux régions, dont une seule est convexe (c'est-à-dire que, si M et N sont deux points quelconques de cette région, le segment MN est tout entier contenu dans cette région) ; cette région est appelée angle saillant xOy. L'autre région est appelée angle rentrant. O est le sommet de l'angle ; les demi-droites Ox, Oy en sont les côtés.



On appelle :

— *Angle plat :* l'un des deux angles définis par deux demi-droites Ox et Oy opposées.

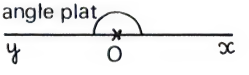
— *Angle nul :* l'angle saillant défini par deux demi-droites confondues.

— *Angle plein :* l'angle rentrant défini par deux demi-droites confondues.

— *Angle droit :* l'angle saillant défini par deux demi-droites perpendiculaires.

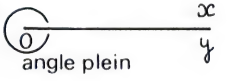
À chaque angle on peut associer un nombre réel, positif ou nul, appelé sa mesure. Il y a trois unités de mesure d'angles : le degré (°), le grade (gr) ou le radian (rd).

Un angle droit a pour mesure 90 degrés, 100 grades ou $\frac{\pi}{2}$ radians ;



Un angle plat — 180 — 200 — π —

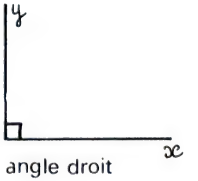
Un angle plein — 360 — 400 — 2π —



Cette définition d'un angle est élémentaire et plutôt dépassée. Actuellement, on préfère définir les angles à l'aide des rotations vectorielles.

Dans le plan orienté, muni d'un repère orthonormé, une rotation de centre O est une application linéaire dont la matrice ◆ est de la forme :

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ avec } a^2 + b^2 = 1.$$



Une rotation transforme un vecteur de longueur 1 en un autre vecteur de même longueur. Inversement, étant donné deux vecteurs $\vec{U}$ et $\vec{V}$ de longueur 1, il existe une rotation de centre O et une seule dans laquelle le vecteur $\vec{U}$ se transforme en vecteur $\vec{V}$. C'est cette rotation qu'on appelle angle orienté des deux vecteurs $\vec{U}$ et $\vec{V}$ et on note cet angle $(\vec{U}, \vec{V})$.

◆ Voir l'article « l'Algèbre linéaire », page 77.

ANNEAU : voir l'article « l'Algèbre », et les pages 46 à 53, 61, 98, 212, 216, 330, 371, 394.

ANTÉCÉDENT : voir Image et les pages 274, 334.

ANTISYMÉTRIQUE page 281.

Prenons l'ensemble $\{ 1, 2, 4 \}$ et choisissons comme relation sur cet ensemble la relation « divise ». On constate que :

1 divise 2, mais 2 ne divise pas 1

1 divise 4, mais 4 ne divise pas 1

2 divise 4, mais 4 ne divise pas 2

Ses seuls couples pour lesquels on a à la fois le premier élément divisant le second et le second divisant le premier sont (1, 1), (2, 2) et (4, 4).

Cette relation est antisymétrique.

Généralisons : soit E un ensemble sur lequel est définie une relation $\mathcal{R}$; cette relation $\mathcal{R}$ est antisymétrique si, quel que soit le couple (x, y), on ne peut avoir à la fois x en relation avec y et y en relation avec x que si x et y sont égaux :

$$\forall x \in E,\ \forall y \in E,\ x\ \mathcal{R}\ y \text{ et } y\ \mathcal{R}\ x \Longrightarrow x = y$$

Cela équivaut à dire que, pour tous les couples (x, y) où x est distinct de y, on ne peut avoir à la fois $x\ \mathcal{R}\ y$ et $y\ \mathcal{R}\ x$ (si l'une des deux phrases est vraie, l'autre ne l'est pas).

APOTHÈME : voir Polyèdres (Pyramide)

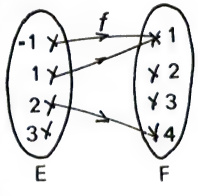
APPARTENIR pages 258 et 259.

Terme primitif de la théorie des ensembles. Synonyme de : « être élément de. »

APPLICATION ◆ pages 17, 209, 231, 277 et 278, 334, 335, 336, 400, 444, 455, 521.

◆ Lire d'abord Relation, Image et Arbre.

Prenons un exemple. La relation « a pour carré » est une application de $E = \{ -1, 1, 2 \}$ dans $F = \{ 1, 2, 3, 4 \}$, puisque chaque élément de E a une image et une seule dans F; mais n'est pas une application de $\{ -1, 1, 2, 3 \}$ dans F puisque 3 n'a pas d'image dans F.



Parmi les relations fonctionnelles d'un ensemble E vers un ensemble F (relations pour lesquelles chaque élément x de E a au plus une image dans F), celles pour lesquelles chaque élément x de E a une image (et une seule) dans F sont appelées applications de E vers F.

Le diagramme sagittal d'une application est tel que de chaque élément de E part une flèche et une seule aboutissant à un élément de F.

Si f est une application de E vers F, on note :

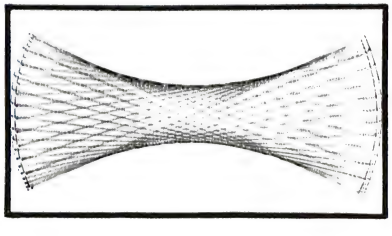
$$f : E \longrightarrow F \quad \text{ou} \quad E \xrightarrow{f} F$$

et, pour exprimer que l'image par f de l'élément x est y, on note :

$$f : x \longmapsto y \quad (\text{ou } f : x \multimap y)$$

Application linéaire : voir l'article « l'Algèbre linéaire ».

APPLICATIONS DES MATHÉMATIQUES : voir l'article pages suivantes et les pages 206, 219.



Les applications des mathématiques

par René Boirel

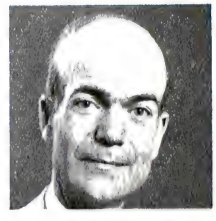
Pourquoi les mathématiques ont-elles tant d'importance de nos jours? Pourquoi leur connaissance est-elle une condition du succès, non seulement dans les examens et concours, mais aussi dans l'activité professionnelle et jusque dans la vie pratique?

C'est que les mathématiques, sciences abstraites d'une rigueur idéale, ont de multiples applications.

Et cela, d'ailleurs, dès l'origine. Leur développement n'a-t-il pas été suscité par des préoccupations d'ordre utilitaire inscrites en filigrane dans les termes mêmes qui désignent les premières mathématiques constituées : la géométrie et le calcul?

En effet, « géométrie », c'est, étymologiquement : « mesure de la Terre », indispensable aux travaux d'arpentage mais aussi aux constructions : pas d'édifices solides sans évaluation précise des aires et des volumes. D'autre part, dans « calcul », n'y a-t-il pas le souvenir phonétique des cailloux utilisés, selon la tradition ancienne, pour effectuer les opérations arithmétiques élémentaires lors des échanges commerciaux? En effet, « calcul » vient du latin « calculus », qui signifie « caillou ».

Dans ces conditions, chaque nouveau développement des mathématiques à partir de ces bases, le calcul et la géométrie, devait, lui aussi, aboutir à des applications pratiques : ce sont les applications classiques des mathématiques.



René Boirel
Docteur ès lettres, professeur à l'Institut national des sciences appliquées de Lyon.
Ses recherches portent sur l'invention dans les techniques et sur la philosophie des sciences.
Il est l'auteur de nombreux ouvrages, notamment *l'Invention* (Paris, P.U.F., 1972), *Théorie générale de l'invention* (Paris, P.U.F., 1961), *l'Art d'inventer* (Paris, Aubanel, 1963).

LES APPLICATIONS CLASSIQUES

Les progrès des mathématiques ont rendu possibles ceux des théories physiques et mécaniques : ne sont-elles pas essentiellement des coordinations mathématiques de résultats de mesures? Sans la géométrie du cercle, Ptolémée n'aurait pu rendre compte des mouvements célestes observés à son époque. De même, Kepler a utilisé

les propriétés de l'ellipse pour imaginer une nouvelle représentation plus précise des trajectoires planétaires. Plus de deux siècles avant Jésus-Christ, en effet, Apollonius de Perge avait étudié les « sections coniques », c'est-à-dire les courbes obtenues en sectionnant un cône de révolution par un plan. Il avait découvert, entre autres, le théorème suivant : « L'aire du parallélogramme construit sur deux demi-» diamètres conjugués est constante et égale au produit des demi-» axes de l'ellipse. » Ce résultat devait permettre à Newton, vingt siècles plus tard, de s'apercevoir que les lois de Kepler expliquaient le mouvement des planètes en termes d'une accélération centrale en r^{-2}, autrement dit selon la loi de l'attraction universelle.

Et que serait l'optique sans la géométrie et sans la trigonométrie ? Si Snellius ◆ n'avait pas eu à sa disposition les notions trigonométriques, il n'aurait pas pu énoncer la loi des sinus. La géométrie est aussi à la base de la mécanique cartésienne, où les « machines simples » s'expliquent « par figures et mouvement ».

◆ Snellius (1580-1626) : astronome et mathématicien hollandais qui mesura la longueur d'un arc de méridien par triangulation et découvrit la loi de la réfraction de la lumière (ou loi des sinus).

Mathématiques et mouvement

Mais c'est surtout la découverte de l'analyse infinitésimale qui a joué un rôle décisif dans le rapide développement de la mécanique, de l'astronomie et des grandes théories de la physique au cours du XVIII^{e} siècle.

En effet, l'esprit humain, pour penser la trajectoire d'un mobile, commence par la décomposer en une série de stations immobiles séparées les unes des autres par des distances aussi petites que possible. Dès lors, si ces distances sont infinitésimales, c'est-à-dire plus petites que toute quantité donnée à l'avance — et telle est justement la définition de l'algorithme différentiel $\mathrm{d}x$ — , il devient possible, en employant les notions du calcul différentiel, de penser d'une manière très précise le mouvement saisi dans sa dynamique même : la dérivée, c'est la vitesse, tandis que l'accélération est la dérivée seconde.

Cette méthode pour mathématiser le mouvement s'applique à l'étude de tout phénomène en évolution. Sous cet angle, la puissance du calcul différentiel est indéniable : si l'on connaît une relation dans l'« infiniment petit évanouissant » qui définit l'état d'un phénomène à l'instant t, on peut penser toute son évolution et, par suite, il devient possible de la contrôler efficacement. En ce sens, écrire l'équation différentielle d'un phénomène, c'est à la fois comprendre sa dynamique et le maîtriser. Quant au calcul intégral, il permet l'évaluation des aires délimitées par une courbe algébrique $f(x)$ et, par conséquent, celle des grandeurs physiques susceptibles d'être représentées par une aire. L'aire comprise entre la courbe représentative d'une fonction $f(x)$ et les parallèles à l'axe des y passant par les points d'abscisses a et b est l'intégrale définie de la différentielle $f(x)\,\mathrm{d}x$ dans l'intervalle (a, b).

Les relations fonctionnelles

D'une manière générale, les mathématiques sont un outil pour le physicien qui cherche à coordonner les résultats d'opérations de mesure : grâce à elles, il met en évidence des relations fonctionnelles entre phénomènes. La plupart des lois physiques ne sont-elles pas de la forme $y = f(x)$? A toute valeur donnée d'une variable x, elles font correspondre une valeur déterminée d'une grandeur mesurable y. La plus simple de ces liaisons fonctionnelles est la *fonction linéaire*, où x et y ne figurent qu'au premier degré, jamais à une puissance supérieure ni non plus au dénominateur : elle exprime une relation de proportionnalité directe entre deux grandeurs. Tel est le cas des allongements subis par un ressort, qui sont proportionnels aux charges qui les produisent, ou encore la proportionnalité entre l'espace parcouru et la durée du mouvement, quand ce dernier est uniforme ($e = vt$).

Quant à la proportionnalité inverse entre deux variables, elle est exprimée par la *fonction homographique*, qui se représente par une hyperbole dont l'équation est $y = \frac{k}{x}$.

Telle est la forme mathématique que revêt notamment la loi de Mariotte : le volume V d'un gaz parfait étant, à température constante, inversement proportionnel à sa pression P, on a la relation $V = \frac{k}{P}$ ou $PV = k$, représentée par une hyperbole équilatère.

Beaucoup de phénomènes physiques ont des lois de cette forme. Par exemple, dans une lentille sphérique, la convergence $\frac{I}{f}$ est inversement proportionnelle au rayon de courbure R ; de même, la capacité d'un condensateur électrique est inversement proportionnelle à la distance e des armatures.

Un autre type de liaison fonctionnelle joue un rôle très important en physique : c'est celui de la *proportionnalité à l'inverse carré*, qui s'écrit $y = \frac{k}{x^2}$. Telle est la relation qui permet notamment d'évaluer la force de gravitation entre deux masses m et m' en fonction de leur distance : $F = k.\frac{mm'}{d^2}$.

Une autre relation est très utilisée : la *proportionnalité au carré*, appelée aussi « fonction quadratique » ou « parabolique », car elle se représente par une parabole dont l'équation peut s'écrire justement : $y = kx^2$.

Sont de cette forme : la loi des espaces parcourus dans la chute des corps, la loi de l'énergie d'un corps en mouvement, qui est fonction

du carré de la vitesse $\left(w = \frac{1}{2} mv^2\right)$, celle de la résistance de l'air qui, pour les vitesses moyennes, est également fonction du carré de la vitesse ($R = kSv^2$).
On rencontre souvent aussi dans les sciences physiques la relation de *proportionnalité à la racine carrée* $y = k\sqrt{x}$: c'est ainsi que la vitesse d'un corps qui tombe est fonction de la racine carrée de sa hauteur de chute ; la période d'un pendule est fonction de la racine carrée de sa longueur.

Quant à la *fonction exponentielle* $y = a^x$, où l'une des grandeurs varie en progression géométrique quand l'autre varie en progression arithmétique, elle exprime une croissance ou une décroissance très rapide. Par exemple, les variations de la pression maximale de la vapeur d'eau en fonction de la température suivent une loi approximativement exponentielle croissante pour les températures élevées, tandis que la décroissance de la pression atmosphérique en fonction de l'altitude suit une loi exponentielle décroissante.
S'il s'agissait d'exprimer la relation inverse, où une grandeur varie en progression arithmétique tandis que l'autre, considérée comme argument, varie en progression géométrique, le physicien utiliserait la *fonction logarithmique* $y = \log_a x$, qui est l'inverse de la fonction exponentielle. Ici, la grandeur considérée comme fonction varie beaucoup plus lentement que la variable.

Pour l'étude des phénomènes périodiques, qui se reproduisent identiquement à eux-mêmes dans le temps, l'usage des *fonctions circulaires* est tout indiqué : en effet, dans ce cas, les grandeurs caractéristiques varient tantôt dans le même sens, tantôt en sens inverse.
Ces phénomènes se représentent donc par des courbes dont la formule comprend des lignes trigonométriques : sinus, cosinus, tangente... La plus simple est la sinusoïde, $y = a \sin x$, qui joue un grand rôle dans les mouvements pendulaires, les phénomènes vibratoires, lumineux ou sonores en particulier, les courants alternatifs, la propagation des ondes hertziennes... A ce propos, on sera peut-être étonné d'apprendre que les « imaginaires », créés au XV^e^ siècle par Tartaglia ◆ et Cardan pour résoudre l'équation du troisième degré et désignés par le symbole i équivalent à $\sqrt{-1}$, permettent de simplifier l'étude des phénomènes périodiques réels : en effet, ces derniers sont représentés par des fonctions trigonométriques, et celles-ci sont liées aux nombres « imaginaires » (notés i), en particulier par la formule d'Euler, $e^{ix} = \cos x + i \sin x$. Il est souvent avantageux de substituer ainsi des exponentielles imaginaires aux lignes trigonométriques exprimant un phénomène ondulatoire : les calculs correspondants s'en trouvent simplifiés.

◆ Niccolo Fontana, dit Tartaglia (1499-1557) : mathématicien italien connu par ses travaux sur l'équation du troisième degré.

Ainsi, ces nombres « imaginaires », employés d'abord avec réticence par les mathématiciens, sont devenus au XIX^e siècle des instruments indispensables pour penser les phénomènes périodiques.
De plus, les vecteurs, qui devaient jouer un si grand rôle dans les applications de l'analyse mathématique à l'étude des forces, ont été introduits en analyse par la théorie des imaginaires. Comment ne pas être étonné devant l'ampleur des conséquences qu'a eues l'invention des « imaginaires » pour l'exploration de la réalité physique? Et cela d'autant plus que, comme le souligne Poincaré, ceux qui les ont créés « ne se doutaient guère du parti qu'on en tirerait pour l'étude » du monde réel : le nom qu'ils leur ont donné le prouve suffi- » samment ◆ ».
Telles sont les principales fonctions et les algorithmes qu'utilisent les sciences physiques.

◆ H. Poincaré : *la Valeur de la science*.

La représentation graphique

Les courbes représentatives de ces fonctions constituent une traduction graphique, qui manifeste nettement le sens de l'évolution des phénomènes dont les états sont coordonnés par ces relations. Dès l'aube des temps modernes, Descartes recommandait l'usage de semblables représentations : supports de l'attention, elles stimulent l'imagination.
Outre ces représentations graphiques de lois fonctionnelles, les sciences physiques ont souvent recours à l'aide précieuse que peuvent leur apporter les diagrammes, qui sont de véritables descriptions des phases d'un phénomène, ou encore les « abaques », ces tableaux graphiques à plusieurs entrées qui comprennent des familles de courbes correspondant à un phénomène et tels qu'une simple lecture remplace toute une série de calculs. Ces diverses représentations graphiques peuvent être rendues encore plus riches de signification par un choix judicieux de coordonnées et d'échelles.

Pour les sciences physiques : un arsenal de relations fonctionnelles

Ce que nous venons de dire est valable également pour la chimie, la biochimie et la physiologie : elles aussi coordonnent des résultats de mesures à l'aide des fonctions mathématiques. En chimie, en biologie comme en physique, les enregistrements graphiques sont utilisés. En particulier, ils font apparaître avec netteté l'amplitude et le rythme de certains phénomènes organiques : tels sont les cardiogrammes, sphygmogrammes, électroencéphalogrammes, précieux auxiliaires de la médecine.
Ainsi, c'est l'ensemble des sciences de la nature qui bénéficie des avantages du symbolisme mathématique.
Le rôle des mathématiques dans les sciences de la nature ne se limite pas à cette fonction de coordination des phénomènes mesurables :

en effet, de ces relations on peut tirer des conséquences qui révèlent des phénomènes non encore observés. Autrement dit, la mathématisation du réel n'est pas uniquement une opération postérieure à l'observation : elle peut la précéder.

L'abstraction mathématique : instrument de découverte expérimentale

Par exemple, l'existence et les caractéristiques de la planète Neptune n'ont-elles pas été déduites par Le Verrier des irrégularités que présentait l'orbite d'Uranus par rapport à l'orbite théorique calculée d'après les lois de Newton? Ces dernières avaient d'ailleurs été découvertes non pas par observation astronomique, mais par l'analyse des relations mathématiques exprimant les lois de Kepler.

Bien plus, des transformations effectuées dans les expressions mathématiques des lois afin de les rendre plus symétriques, plus harmonieuses, c'est-à-dire plus « belles » pour le mathématicien, peuvent recevoir, par la suite, une confirmation expérimentale. Dans ce cas encore, les mathématiques devancent d'une manière étonnante l'expérience.
C'est ce qui s'est produit notamment quand Maxwell, poussé par son goût pour les formules symétriques, eut l'audace d'ajouter aux équations de l'électrodynamique un terme qui, trop petit pour produire des effets appréciables avec les méthodes anciennes, devait recevoir une signification expérimentale vingt ans plus tard.

La mathématisation du réel révèle des analogies de structure

A ce niveau, la mathématisation du réel permet non seulement de devancer l'expérience, mais encore d'apercevoir des analogies de structure entre des phénomènes en apparence disparates, qu'on n'aurait pas soupçonnées sans ce révélateur que constituent les similitudes entre formules mathématiques : Maxwell n'est-il pas parvenu ainsi à unifier dans une vaste construction mathématique les phénomènes électromagnétiques et les phénomènes lumineux?

Ce triple rôle que jouent les mathématiques dans les sciences de la nature — coordination des mesures, instrument de découverte, révélateur de structures communes — permet d'affirmer que la science est bien, selon la formule de Bergson, « fille des mathématiques ».

Technique et mathématique

A fortiori, il en est de même pour les techniques à fondement scientifique. C'est pourquoi les mathématiques occupent une place centrale dans la formation d'un ingénieur.

Sans géométrie descriptive, pas de plans précis. La mécanique rationnelle, le calcul différentiel et intégral guident l'édification des barrages. Et, sans nombres complexes, que ferait l'électrotechnicien ? En effet, nous avons vu que les « imaginaires » permettent des simplifications heureuses dans les calculs relatifs aux phénomènes ondulatoires et, par suite, dans ceux concernant les courants alternatifs.

Si la réalisation de la première machine Gramme ne nécessitait pas des connaissances mathématiques très approfondies, par contre, ce n'est qu'un ingénieur à haute formation mathématique, comme Tesla, qui pouvait concevoir le moteur asynchrone à champ tournant.

La théorie des fonctions analytiques de Cauchy ◆ est à la base de l'aérodynamique : en ce sens, elle conditionne la construction aéronautique. L'idée des profils « laminaires », dont la résistance à l'avancement est minime, n'a-t-elle pas été suggérée par des développements mathématiques très abstraits ?

◆ A.L. Cauchy (1789-1857) : mathématicien français, dont les travaux extrêmement nombreux portent surtout sur les notions de différentielles, d'intégrales, de continuité, etc.

Comme les sciences physiques, les recherches techniques font usage des graphiques et surtout d'abaques, à tel point qu'une science des abaques, la nomographie, a été constituée par Maurice d'Ocagne pour rendre encore plus puissant cet instrument de représentation des phénomènes intervenant dans les dispositifs techniques.

Les applications des mathématiques stimulent leur progrès

Ces applications classiques des mathématiques ont eu des répercussions sur la recherche mathématique elle-même.

La théorie des équations différentielles et aux dérivées partielles ainsi que le calcul des variations se sont développés au XVIII[e] siècle sous l'impulsion des problèmes que posait l'expression mathématique des phénomènes astronomiques et physiques étudiés à cette époque. Les travaux de Riccati ◆, Euler ◆, Clairaut ◆, Legendre ◆ sur les équations différentielles sont inséparables de problèmes de mécanique. L'étude mathématique des phénomènes de conductibilité de la chaleur dans un conducteur a conduit Fourier à entreprendre l'approfondissement systématique des fonctions discontinues non développables en série de Taylor. Pour établir les équations thermologiques, il a dû forger un nouvel instrument mathématique, les fameuses séries qui portent son nom et qui devaient bouleverser l'analyse mathématique.

◆ Riccati (1676-1754) : mathématicien italien connu par ses travaux sur les équations différentielles.

◆ L. Euler (1707-1783) : un des plus grands hommes de science suisses, qui a su tirer parti d'un grand nombre de théories antérieures grâce à son génie d'unification et de systématisation.

◆ Clairaut (1713-1765) : mathématicien français connu par ses travaux sur les courbes, la mesure d'un degré du méridien et la prédiction du retour d'une comète.

De même, des ingénieurs ont été amenés à créer les outils mathématiques dont ils avaient besoin pour traiter certains problèmes techniques. C'est ainsi qu'a été conçue l' « équation des télégraphistes » pour exprimer la déformation du courant dans les câbles transatlantiques ; elle a été résolue par W. Thomson (Lord Kelvin). C'est également pour faciliter les calculs des électrotechniciens que Heaviside ◆ a eu l'idée du « calcul symbolique ». Les bureaux d'études et les laboratoires l'ont utilisé avant qu'il ait droit de cité en mathématiques pures, ce qui n'a été possible qu'à la suite de travaux contemporains qui l'ont fondé rigoureusement.

◆ Legendre (1752-1833) : mathématicien français connu par ses travaux sur la trigonométrie et la théorie des nombres.

◆ Heaviside (1850-1925) : mathématicien et physicien anglais qui a créé le « calcul symbolique ».

Donc, si les mathématiques imprègnent la civilisation technicienne, où convergent leurs applications classiques, on peut dire aussi que leur développement subit le contrecoup des problèmes qui se posent dans les sciences physiques ou dans l'industrie.

LES APPLICATIONS ACTUELLES

L'emprise des mathématiques sur les sociétés industrielles modernes s'étend encore plus loin, jusqu'à l'organisation non seulement de l'activité économique, mais encore de la vie sociale par l'intermédiaire des sciences humaines, qui les utilisent de plus en plus. Mais, d'abord, leur importance pour les sciences théoriques comme pour les applications techniques s'est encore accrue : sans des développements relativement récents des mathématiques, la science contemporaine ne serait pas ce qu'elle est. Les théoriciens ont puisé dans l'arsenal mathématique des notions qui datent de moins d'un siècle ou qui ont été approfondies depuis peu.

Probabilités et statistiques

Les modèles particulaires de la matière font appel aux statistiques et aux probabilités. S'il est vrai que les phénomènes matériels sont formés de particules en mouvement qui s'entrechoquent ou interagissent, leurs propriétés observables doivent être, en effet, la manifestation globable, statistique, des mouvements et chocs de ces particules.
Selon cette perspective, qui a été adoptée par les fondateurs des théories cinétiques, les propriétés observées sont le résultat des mouvements incoordonnés de ces innombrables particules : dès lors, c'est le calcul des probabilités qui est à la base de leur interprétation rationnelle.

Probabilités et statistiques : indispensables aujourd'hui pour le physicien des particules

Ainsi, cette discipline mathématique, qui s'était constituée aux XVII^{e} et XVIII^{e} siècles, un peu en marge des recherches classiques, sous la forme d'une théorie des jeux de hasard, a vu s'élargir considérablement son champ d'application avec les conceptions cinétiques : les notions de pression, de température, de chaleur spécifique, etc., s'éclairent quand on les envisage comme résultant du mouvement et des chocs des particules, et cela quel que soit l'état de la matière considérée, qu'il s'agisse de gaz, de liquides ou de solides. Dès lors, les probabilités et les statistiques sont un fil d'Ariane précieux pour explorer les aspects corpusculaires de la matière.
Cette expansion des théories cinétiques de la matière a mis en évidence des problèmes inédits, qui ont motivé un nouveau développement du

calcul des probabilités. Par exemple, l'étude du mouvement brownien, constitué par une agitation chaotique de particules, n'a-t-elle pas suscité les travaux aboutissant à l'équation de Capmann-Kolmogoroff ◆ et à la conception des « probabilités en chaîne »?

◆ Kolmogoroff : mathématicien russe né en 1903, fondateur de la théorie axiomatique des probabilités.

Le calcul vectoriel

Quant à la théorie des champs et aux conceptions relativistes, qui caractérisent aussi la physique contemporaine, elles reposent sur le calcul tensoriel et les espaces riemanniens. Déjà, les travaux d'hydrodynamique concernant le mouvement des fluides et la théorie de l'élasticité élaborée pour interpréter le mouvement interne des solides avaient mis en évidence, dès le début du XIX^{e} siècle, certaines structures mathématiques dont l'ensemble une fois codifié devait constituer plus tard le calcul vectoriel, qui applique les opérations algébriques aux grandeurs orientées.

C'est surtout la théorie de Maxwell, unifiant les phénomènes électromagnétiques dans une conception fondée sur les champs et leur propagation de proche en proche, qui a habitué les physiciens à utiliser le calcul vectoriel. Or celui-ci devait être généralisé par le calcul tensoriel ; il s'est produit pour ce dernier calcul un peu ce qui a eu lieu au XVIII^{e} siècle pour les équations différentielles à la suite de leur application à des problèmes concrets : la relativité ayant révélé aux physiciens les possibilités de ce calcul, ils l'ont appliqué aux diverses branches de la physique, ce qui a posé des problèmes stimulants pour les mathématiciens ; ceux-ci lui ont alors donné des bases rigoureuses.

Ce nouveau symbolisme employé systématiquement, jusque dans les domaines physiques considérés comme classiques, a mis en évidence des relations et des analogies insoupçonnées jusqu'alors entre des phénomènes en apparence hétérogènes.

Il en est de même pour la conception des espaces polydimensionnels : sans Gauss ◆, qui, en introduisant l'usage des coordonnées curvilignes, a jeté les bases de la théorie générale des surfaces, sans la théorie des espaces de Riemann ◆ généralisant l'emploi des coordonnées curvilignes et fondant rigoureusement la métrique d'un espace continu à un nombre quelconque de dimensions, sans les travaux de Levi-Civita ◆ aussi, Einstein n'aurait pu élaborer la relativité « généralisée », ainsi appelée parce qu'on y tient compte des mouvements accélérés et du caractère absolu des accélérations.

◆ K.F. Gauss (1777-1855) : mathématicien et astronome allemand, qui fut à l'origine de travaux essentiels sur le calcul des probabilités.

◆ B. Riemann (1826-1866) : mathématicien allemand, surtout connu comme auteur d'une géométrie non euclidienne.

◆ Levi-Civita (1873-1941) : mathématicien italien qui a créé le calcul différentiel absolu.

La relativité généralisée : un stimulant pour les mathématiciens

Selon cette théorie, en effet, une masse matérielle donne aux régions voisines de l'espace-temps une courbure, de sorte que localement l'espace-temps est une surface courbe à quatre dimensions.

Cela a permis d'expliquer l'anomalie observée dans l'avance du périhélie ◆ de Mercure par rapport aux prévisions newtoniennes : la masse du Soleil courbant l'espace-temps E autour de lui, les trajectoires spatio-temporelles des planètes sont des « géodésiques », ou lignes courbes du plus court chemin entre deux points, tracées dans E. En faisant les calculs correspondants, on trouve effectivement une avance de 42″ par siècle du périhélie de Mercure.
La relativité généralisée a habitué les physiciens théoriciens à utiliser les espaces polydimensionnels ; et ces applications, en montrant leur intérêt pratique, ont incité des mathématiciens tels Elie Cartan et Hermann Weyl à en approfondir l'étude.

◆ Périhélie : point où l'orbite d'une planète est le plus proche du Soleil.

Calcul matriciel et équations intégrales

Quant à la physique des quanta, qui est peut-être l'aspect le plus étonnant de la science contemporaine, elle doit son développement à d'autres théories mathématiques de création récente : le calcul matriciel et les équations intégrales. Certes, au début, cette conception discontinue du mouvement des particules et du rayonnement ne faisait appel qu'à des notions mathématiques classiques, bien connues des physiciens. Mais tout changea quand N. Bohr ◆ introduisit dans la théorie des quanta le principe de correspondance, qui nécessitait l'emploi de certaines théories de mécanique analytique, généralement ignorées des physiciens.
Avec la mécanique quantique de W. Heisenberg ◆, qui repose sur les matrices algébriques, ce sont des notions mathématiques de création plus récente, motivées par l'étude des transformations linéaires, qui deviennent un instrument indispensable en physique théorique.
Le calcul matriciel y joue un rôle de plus en plus grand. Ne'eman ◆ a même utilisé des matrices pour représenter d'une manière commode les interactions entre particules atomiques. Gell-Mann ◆ devait tirer de ces considérations mathématiques l'hypothèse des « quarks », ces particules à partir desquelles toutes les autres pourraient être construites.
Comme pour le calcul tensoriel, les applications des matrices à la physique ont suscité le développement du calcul matriciel. Notamment, on a été amené à forger la notion de dérivée d'une matrice par rapport à une autre matrice.
Ces progrès du calcul matriciel ont étendu son emprise bien en dehors de la physique des quanta, jusqu'en électrotechnique, où il permet de résoudre élégamment les problèmes relatifs aux quadripôles ou aux filtres électriques, et jusque dans l'industrie : utilisé d'abord en aéronautique, il a envahi progressivement les bureaux d'études.
Le développement de la physique des quanta a été conditionné aussi par celui de la théorie mathématique des équations intégrales, due

◆ Niels Bohr (1885-1962) : physicien danois, prix Nobel en 1922, qui a conçu un modèle de l'atome, une théorie des transmutations nucléaires et le principe selon lequel corpuscules et ondes représentent deux aspects « complémentaires » d'une même réalité.

◆ Werner Heisenberg : physicien allemand né en 1901, prix Nobel en 1932. Spécialiste de mécanique quantique, il a établi les « relations d'incertitude » (il est impossible d'attribuer à une particule à la fois une position et une vitesse déterminées).

◆ Ne'eman : physicien israélien qui est avec Gell-Mann à l'origine d'une classification des particules appelée « modèle octuple ».

◆ Gell-Mann : physicien américain, né en 1929, spécialiste de la physique des particules.

surtout aux travaux de Fredholm ◆. Celle-ci fournit de puissants moyens d'investigation théorique applicables à divers domaines : leur perfectionnement est stimulé par les problèmes que la nouvelle physique des quanta pose au mathématicien.

◆ Fredholm (1866-1927) : mathématicien suédois connu par ses travaux sur les équations intégrales.

La théorie des groupes

Cette branche des mathématiques, de création relativement récente, est devenue l'auxiliaire de la mécanique ondulatoire, qui représente l'état des particules « élémentaires » par certaines fonctions d'onde. Le calcul de ces fonctions pour un système atomique à plusieurs constituants devient rapidement inextricable quand le nombre de ces constituants croît. Grâce à la théorie des groupes, des considérations de symétrie permettent de trouver certaines caractéristiques globales des fonctions d'onde, d'où l'on peut déduire les propriétés du système étudié, sans avoir à connaître exactement les fonctions d'onde.

Cette branche très abstraite des mathématiques, qui les unifie par certaines structures fondamentales de transformation qu'on retrouve aussi bien en analyse qu'en géométrie, a donc des applications pratiques pour la physique atomique : sans elle, on ne pourrait pas prévoir rapidement les comportements des atomes complexes. En ce sens, les industries de l'atome dépendent de la théorie des groupes. C'est également sur la théorie des groupes qu'est fondée la cristallographie moderne : elle permet de comprendre la raison profonde de l'organisation des cristaux ◆. Quand on pense à l'importance de ces derniers dans la matière solidifiée, on reste stupéfait devant la puissance d'une théorie purement mathématique pour éclairer le réel.

◆ Cristal : solide naturel limité par des faces planes qui résulte d'un assemblage ordonné d'atomes (c'est le contraire d'un corps « amorphe », où les atomes sont disposés sans ordre).

De même, sans la très abstraite théorie des fonctions et valeurs propres des équations différentielles aux dérivées partielles, qui démontre l'existence des états stationnaires relatifs notamment aux phénomènes vibratoires et détermine la forme des fonctions qui les représentent, les spécialistes des télécommunications ne pourraient pas étudier avec précision la propagation des ondes électromagnétiques très courtes à l'intérieur de tubes métalliques dits « guides diélectriques », pour lesquels ils emploient aussi les fonctions de Bessel ◆, de Laplace, les polynômes de Legendre.

◆ Bessel (1784-1846) : astronome allemand.

On se rend compte à quel point les abstractions de l'analyse mathématique sont devenues aujourd'hui des moyens au service du progrès non seulement théorique, mais aussi technique. Ainsi, les résultats mathématiques en apparence les plus éloignés du réel sont utilisés actuellement dans les laboratoires, pour la recherche fondamentale comme pour la recherche appliquée.

C'est pourquoi on a constaté, au cours des cinquante dernières années, un décalage croissant entre l'enseignement classique des mathématiques, qui préparait traditionnellement aux études scientifiques ou techniques supérieures, et l'outillage mathématique dont

avaient besoin effectivement les chercheurs. Une réforme des programmes est devenue nécessaire : il a fallu y introduire des théories mathématiques réservées jusqu'alors à des spécialistes et qui sont maintenant indispensables pour suivre le progrès scientifique.
C'est ainsi que l'on a vu apparaître, d'abord dans les classes préparatoires aux grandes écoles puis dans les classes terminales des lycées, les « mathématiques modernes », qui envahissent actuellement tout l'enseignement secondaire et sont même enseignées dès l'école primaire.

LES APPLICATIONS DES MATHEMATIQUES MODERNES

Parmi les nouvelles mathématiques enseignées actuellement, il convient de distinguer, d'une part, ce qu'on pourrait appeler les « mathématiques à applications modernes », mais de création plus ancienne, telles que le calcul vectoriel, les matrices, le calcul des probabilités, l'algèbre de Boole, l'analyse combinatoire, la topologie, et, d'autre part, les « mathématiques modernes » à proprement parler, qui sont une présentation axiomatique des diverses branches des mathématiques unifiées autour de la théorie des ensembles. C'est pourquoi on désigne justement celles-ci par l'expression « la mathématique moderne ». Sous sa forme bourbakiste ◆ du moins, cette axiomatisation des mathématiques date de quelques décennies à peine.

◆ De N. Bourbaki, fictif mathématicien suédois. Ce terme désigne en fait un groupe de jeunes mathématiciens français contemporains, qui se sont donné pour tâche, vers les années 1930, de repenser les fondements des mathématiques suivant la méthode axiomatique.

LES « MATHÉMATIQUES A APPLICATIONS MODERNES »

Considérons d'abord les « mathématiques à applications modernes ». Nous avons dit l'importance du calcul vectoriel, des matrices, des statistiques et probabilités pour la physique contemporaine. Or ces branches des mathématiques ont des applications actuelles non seulement dans les sciences de la nature, mais encore dans les sciences humaines.

En économie, en politique

Imagine-t-on aujourd'hui qu'un économiste, un sociologue ou un psychologue puisse ignorer les statistiques et le calcul des probabilités ? Ne serait-ce que pour bien comprendre la signification des résultats élaborés à l'aide de ces instruments pour traiter les informations, leur connaissance est indispensable.
Le jour où le chevalier de Méré, après une nuit passée à jouer aux dés,

vint demander à Pascal : « En combien de coups peut-on espérer » faire sonner [c'est-à-dire amener deux 6] avec deux dés? », il ne se doutait pas qu'il allait être à l'origine d'une nouvelle discipline mathématique appelée à bouleverser, trois siècles plus tard, pour l'homme d'action, l'approche de la réalité sociale et la préparation des décisions. Il aurait été bien étonné si on lui avait annoncé qu'un mathématicien français, Augustin Cournot ◆, préconiserait en 1838, dans ses « Recherches sur les principes mathématiques de la théorie des richesses », l'utilisation du calcul des probabilités dans le domaine économique, et qu'il inviterait les polytechniciens de l'époque à perfectionner ses modèles de concurrence.

◆ A. Cournot (1801-1877) : mathématicien, économiste et philosophe français.

Les travaux de Cournot ont été développés à la fin du XIX^{e} siècle par les économistes de l'école suisse, Walras ◆ et Pareto ◆. Plus récemment, l'économétrie américaine a révélé leur portée : la théorie des jeux s'est transformée en une théorie mathématique des comportements économiques. Telle est la raison qui a motivé, tardivement, l'introduction des probabilités dans les programmes d'économie politique.

◆ Walras (1834-1910) : économiste français, créateur de l'économétrie mathématique.

◆ Pareto (1848-1923) : économiste italien, successeur de Walras dans l'exposition mathématique de l'économie.

Des travaux de Condorcet aux sondages d'opinion

De même, les travaux mathématiques de Condorcet ◆ sur les décisions obtenues par vote sont restés longtemps inaperçus : il a fallu presque un siècle et demi pour qu'on s'aperçoive de la profondeur de cette œuvre. Aujourd'hui, à la suite des récentes consultations électorales, nul ne conteste l'efficacité du calcul des probabilités et des statistiques pour prévoir l'orientation globale d'un vote à l'aide des sondages d'opinion en fonction d'un échantillonnage judicieux et surtout pour estimer dans une « fourchette » très limitée les divers pourcentages obtenus, à partir des premiers résultats partiels après la clôture du scrutin.

◆ Condorcet (1743-1794) : philosophe et mathématicien français. Conventionnel, il élabora un plan d'organisation de l'Instruction publique.

L'aide à la décision

Le calcul des probabilités est également utile pour la préparation des décisions dont dépendent les réalisations qui s'effectuent progressivement dans le temps : le succès repose alors, en effet, sur une juste appréciation de l'évolution future des situations présentes.

L'industriel, l'économiste, le politique ou, plus généralement, le stratège dans tous les domaines doit choisir entre plusieurs manières d'agir en présence de conjonctures en cours d'évolution : il ne peut donc pas connaître avec certitude ce que ces dernières seront effectivement devenues quand ses décisions auront des conséquences. Dans ces conditions, la réussite d'une stratégie repose sur une estimation de probabilités.

Descartes se souvient manifestement de pareilles expériences, rencontrées dans sa vie de mousquetaire, quand il reconnaît que la règle d'évidence — on ne doit fonder le jugement que sur des certi-

tudes et non sur des probabilités — qu'il pose comme principe directeur de la vie intellectuelle est impraticable dans l'action : cette dernière « ne souffrant souvent aucun délai, c'est une vérité » très certaine, affirme-t-il, que lorsqu'il n'est pas en notre pouvoir » de discerner les plus vraies opinions nous devons suivre les plus » probables ◆ ».

◆ R. Descartes : *Discours de la méthode* (deuxième règle de la morale provisoire).

Résoudre un problème de décision, c'est en effet prendre le risque de choisir entre des solutions, où la part du probable peut être réduite — sans qu'on puisse l'éliminer entièrement — au profit des certitudes. Cette situation inconfortable, que connaît bien l'homme d'action, n'a-t-elle pas été symbolisée par le jeu des échecs, où chaque coup modifie la partie, dont l'évolution dépend aussi des décisions ultérieures de l'adversaire, qui ne peuvent être que supposées?

Pour évaluer d'une manière plus précise ces pyramides de probabilités sur lesquelles repose le succès d'une entreprise, autrement dit pour clarifier les données motivant une décision, des instruments mathématiques ont été conçus et développés récemment : leur ensemble constitue ce qu'on appelle la « recherche opérationnelle », parce que ces travaux ont été stimulés par la préparation de certaines « opérations » militaires pendant la Seconde Guerre mondiale.

Des probabilités à la recherche opérationnelle

Il s'agit d'une mathématisation des facteurs essentiels qui entrent en jeu dans les problèmes d'organisation militaire, économique, industrielle. Elle utilise non seulement le calcul des probabilités, mais aussi des branches récentes des mathématiques, telles que la théorie des graphes : ces ensembles de points, joints entre eux par des lignes ou des flèches symbolisant une certaine relation, peuvent représenter en effet des localités, des ateliers, des entrepôts, les étapes d'une production... Notamment, la méthode d'organisation appelée le P.E.R.T. (Process Évaluation and Review Tasks) est devenue en quelques années un puissant moyen pour planifier les diverses tâches que comporte la réalisation d'un projet important : celles-ci sont pensées clairement dans leurs relations mutuelles à l'aide de graphes. Par exemple, un succès astronautique repose sur la coordination préalable d'activités multiples. Cette méthode fournit, en quelque sorte, une carte des événements : elle indique les étapes à parcourir ; surtout, elle permet d'en contrôler méthodiquement la réalisation. La recherche opérationnelle a mis au point également des modèles mathématiques pour rendre plus rationnelle la gestion des stocks, où il faut tenir compte à la fois de la demande d'articles, du réapprovisionnement nécessaire et des coûts de stockage, de lancement et de pénurie. De même, les problèmes que posent la dépréciation, l'usure et le renouvellement des équipements peuvent être pensés et clarifiés par ces méthodes, qui permettent d'élaborer des modèles mathématiques pour les phénomènes économiques mis en jeu par les investissements correspondants.

Également, les problèmes d'optimisation motivés par le souci d'améliorer le rendement d'une entreprise peuvent être traités de cette manière. Déjà, en 1776, Monge, alors professeur à l'école du Génie de Mézières, n'étudiait-il pas mathématiquement la façon la moins onéreuse d'organiser les travaux de déblais et remblais ?

Toutes ces méthodes mathématiques du traitement des informations économiques ou sociales reçoivent aujourd'hui une aide considérable des ordinateurs électroniques.

L'ordinateur et les mathématiques

Or ces derniers ont mis en honneur certaines branches des mathématiques qui, jusqu'alors, n'étaient cultivées que par des spécialistes : notamment la numération binaire et la logique de Boole. Du coup, leur connaissance devient indispensable pour tous ceux qui ont à utiliser les ordinateurs. C'est la raison pour laquelle les enfants sont initiés maintenant, dès l'école primaire, à la numération binaire : nous sommes entrés dans l'ère de l'« informatique », c'est-à-dire du traitement des informations par machines.

D'une curiosité mathématique...

La numération binaire, qui était, il y a peu de temps encore, une sorte de curiosité mathématique, est devenue l'instrument par excellence permettant de traduire dans le système des signaux électriques ou électroniques les expressions numériques : en effet, la numération binaire ne comprend que les chiffres 0 et 1, auxquels on peut faire correspondre respectivement la fermeture et l'ouverture d'un circuit électrique ou électronique. Dans un appareillage de ce type, le passage du courant exprime techniquement le chiffre 1, tandis que son interruption est l'équivalent technique du chiffre 0. En un sens, les ordinateurs sont essentiellement des systèmes complexes d'interrupteurs qui fonctionnent des milliers de fois par seconde et sont, par suite, susceptibles de traduire très rapidement en termes d'ouverture et de fermeture des circuits tous les nombres et les opérations faites sur eux. Ils peuvent donc effectuer ces dernières en un temps record.

...à une logique du tout ou rien

Le courant passe ou ne passe pas : les ordinateurs incarnent dans leurs circuits la logique de Boole◆. Celle-ci, qui, à l'époque de sa création, au XIX^{e} siècle, n'avait intéressé personne, sinon quelques rares spécialistes, tant cette théorie très abstraite des propositions logiques semblait alors pure spéculation, est ainsi mise au premier plan de l'actualité par l'invention des ordinateurs. Elle est à la base même de leur construction et, par extension, de l'informatique.
Les performances des ordinateurs ont amplifié considérablement les applications du calcul des probabilités et des méthodes statistiques.

◆ G. Boole (1815-1864) : logicien et mathématicien anglais, célèbre pour ses théories sur le calcul et son œuvre en logique symbolique.

L'ordinateur et son arsenal mathématique ont fait leur entrée dans la gestion des entreprises : c'est pourquoi les mathématiques ont une telle importance à H.E.C., à E.S.S.E.C., comme dans les diverses écoles supérieures de commerce.

L'ordinateur et les sciences de l'homme

On a même construit des machines pour résoudre les « problèmes de décision » de l'économiste. Il vaudrait mieux dire d'ailleurs « machines destinées à résoudre les problèmes de probabilité dont dépendent les choix entre plusieurs stratégies ». En effet, si elles fournissent des indications utiles à l'homme d'action, elles ne le déchargent pas de ses responsabilités et ne le dispensent aucunement de prendre une décision : il est certes mieux informé par ces ordinateurs, mais il lui reste en définitive à apprécier la valeur respective des diverses solutions proposées et à en choisir une.
En somme, ces machines « à décider » éclairent le choix sans l'imposer : par l'analyse des conditions dont l'enchevêtrement complexe commande le succès d'une organisation industrielle, commerciale ou économique, elles dégagent les solutions qui ont le plus de chances d'être efficaces, vu la conjoncture et son évolution probable, sans désigner pour autant celle qui sera effectivement la meilleure, car la réalité du futur reste inconnue, même s'il est entrevu comme possibilité.
Il en est de même des machines pour élaborer un diagnostic : en effet, l'ordinateur et ses mathématiques ont fait récemment leur entrée aussi dans les sciences de l'homme, notamment en médecine. Puisqu'un diagnostic repose essentiellement sur la comparaison des signes recueillis chez le malade avec les symptômes décrits dans les traités de pathologie, il peut être considéré comme l'aboutissement d'un raisonnement rigoureux. Dans ces conditions, on comprend qu'une rationalisation intégrale du diagnostic médical ait pu tenter les mathématiciens. Notamment, un médecin passionné de mathématiques, M. F. Paycha, aidé par un ingénieur du Génie maritime, M. Abraham, a inventé des formules permettant d'apprécier la probabilité d'un diagnostic en fonction de la qualité et du nombre des signes observés. Dès lors, la réalisation d'une machine programmée en vue de la mise en action de formules semblables devenait possible. Mais il faut bien se rendre compte qu'un tel ordinateur ne résout que les problèmes de probabilité surgissant à l'occasion du diagnostic à établir, en fonction des informations qu'on lui aura fournies. Or ces dernières n'ont pas une signification en elles-mêmes, considérées isolément : leur signification réelle dépend de leur intégration dans l'ensemble des symptômes. Cette connaissance du malade perçu comme un tout incombe au médecin. Mais l'ordinateur peut, en éclairant son jugement, l'aider considérablement pour son diagnostic.

Dans les sciences de l'homme, ce n'est pas seulement la connaissance de l'organisme humain qui relève maintenant de l'ordinateur : c'est l'ensemble de la recherche en sciences humaines qui lui est redevable de ses succès actuels.

La linguistique, notamment, ne peut s'en passer : les statistiques sur l'emploi des mots et tournures dans un groupe de textes à telle époque mettent en évidence les structures d'une langue à une phase de son évolution. Ce sont les statistiques faites avec l'aide d'ordinateurs qui ont permis le renouvellement de l'analyse des textes. Mieux que toute autre méthode, elles révèlent les emprunts, les textes apocryphes, voire les auteurs multiples cachés sous un nom unique. Le progrès de la formalisation mathématique dans des domaines considérés jusqu'à ces dernières années comme réservés aux littéraires n'est-il pas un signe des temps?

Toute discipline qui recueille et interprète des informations relève des techniques mathématiques de traitement des informations, autrement dit de l'informatique.

En particulier, que seraient aujourd'hui la sociologie, la démographie sans statistiques et privées d'ordinateurs? C'est pourquoi des cours de mathématiques ont lieu à l'heure actuelle dans les U.E.R. de lettres et sciences humaines.

L'ordinateur rend d'autres services aux disciplines qui traitent des informations : il combine rapidement ces dernières et permet d'apercevoir toutes les variantes d'une situation et ses divers développements possibles.

Il rend manifeste la puissance des méthodes combinatoires. Raymond Lulle ◆ en avait posé les bases dès le XIII^{e} siècle dans son « Grand Art », où il décrit une véritable machine logique pour former de nouvelles propositions par combinaison ordonnée de signes inscrits sur des cercles concentriques mobiles les uns par rapport aux autres. Mais c'est surtout Leibniz qui devait donner une armature mathématique à ce projet par l'analyse combinatoire. Dans son « Ars combinatoria », il prophétisait qu'elle guiderait dans le futur l'étude, à l'aide de machines logiques, des multiples possibilités découlant d'une situation en cours de variation ; l'ordinateur permet la réalisation de ce projet leibnizien.

◆ Raymond Lulle (1233-1315) : théologien et logicien espagnol, mystique et homme d'action, précurseur, avec son *Grand Art*, du calcul logique et des machines logiques.

Sous cet angle, l'informatique utilise un autre type de mathématiques « modernes », qui découle certes des travaux de Leibniz, mais qui n'a reçu sa pleine expression que dans le cadre de la théorie des ensembles. En ce sens, l'informatique est fille de « la mathématique moderne » autant que des « mathématiques à applications modernes ».

LES APPLICATIONS DE « LA MATHÉMATIQUE MODERNE »

Le rêve leibnizien de combinatoire universelle est devenu aujourd'hui réalité par la médiation de la théorie des ensembles. En effet, cette

dernière envisage des systèmes d'éléments — ou « ensembles » — caractérisés principalement par les relations qu'ont entre eux les éléments. L'analyse combinatoire consiste essentiellement à faire varier la position des éléments constituants les uns par rapport aux autres d'une manière ordonnée : elle trouve donc dans la théorie des ensembles un cadre approprié qui lui donne une portée plus générale, car, dans cette perspective, on ne s'occupe que des relations entre les éléments, indépendamment de la nature propre de ces derniers.

C'est justement pour cette raison que la théorie des ensembles unifie les différentes branches de l'analyse et de la géométrie autour de structures fondamentales, qui se retrouvent identiques quelle que soit la nature des êtres mathématiques considérés et d'où découlent leurs propriétés respectives : ces structures rendent compte, en définitive, de leurs transformations possibles et des constructions qu'on peut organiser à partir d'eux. Dès lors, on comprend qu'avec la théorie des ensembles les mathématiques soient devenues « la mathématique ».

Dans la mathématique moderne, l'accent est mis sur les relations

Dans cette nouvelle perspective, les relations constituent ou unissent les « objets » qui étaient traditionnellement étudiés pour eux-mêmes : droites, cercles, angles, nombres, fonctions... Elles étaient auparavant secondes par rapport aux êtres mathématiques : comme eux, elles étaient de nature numérique, géométrique, fonctionnelle... Avec « la mathématique », au contraire, ce sont les relations considérées en elles-mêmes, étudiées pour elles-mêmes, qui ont la priorité.

Une théorie des relations et des opérations s'est donc constituée avec un symbolisme précis, dont le champ d'application est multiple. Comme il est dit dans Bourbaki : « C'est sans doute la possibilité » de ces extensions successives dans lesquelles la forme des calculs » restait la même alors que la nature des êtres mathématiques soumis » à ces calculs variait considérablement qui a permis de dégager » peu à peu le principe directeur des mathématiques modernes, à » savoir que les êtres mathématiques pris en eux-mêmes importent » peu ; ce qui compte, ce sont leurs relations. ◆ »

◆ N. Bourbaki : in *Algèbre*, introduction.

Dans une telle perspective, les considérations d'ordre — en particulier entre des éléments constituants — deviennent primordiales. C'est pourquoi la théorie des ensembles étend à toutes les branches des mathématiques classiques et, plus généralement, à toute structure les opérations de l'analyse combinatoire, qui font varier la disposition des éléments les uns par rapport aux autres d'une manière ordonnée.

On conçoit alors que la connaissance de « la mathématique » soit indispensable à ceux qui vont vivre à l'ère de l'informatique : c'est le cas notamment des jeunes générations.

Par sa puissance combinatoire, l'ordinateur peut non seulement

étendre l'aperception des possibles, qui demeure souvent limitée pour l'intelligence humaine, mais encore stimuler par ce moyen l'imagination. En ce sens, l'informatique et, par suite, la mathématique moderne qui en est le fondement sont susceptibles de contribuer au développement de la créativité.
C'est ce que confirment certaines expériences toutes récentes faites dans des laboratoires de recherche, où l'on a fait « dialoguer » chercheurs et ordinateurs pour élaborer des hypothèses de travail. D'ailleurs, sans faire usage d'ordinateurs, le seul emploi du symbolisme de la mathématique moderne facilite le déploiement de la pensée combinatoire et, par suite, de l'imagination créatrice.
Mais ce n'est pas uniquement dans l'organisation de la vie économique ou la recherche scientifique et technique que les symboles de la mathématique moderne se révèlent féconds. Ils présentent un grand intérêt aussi pour les sciences humaines.

Les ensembles et les sciences humaines

En effet, jusqu'à présent, les applications des mathématiques classiques dans ce domaine restaient limitées principalement à des enquêtes statistiques, d'ailleurs utilisées souvent avec réticence car les faits sociaux se réduisent difficilement, et toujours imparfaitement, à du quantitatif.

Le quantitatif passe au second plan...

Or, une opération se définit dans la mathématique moderne en dehors de toute référence quantitative : n'apparaît-elle pas, dans la perspective des ensembles, comme la mise en correspondance d'un élément appartenant à l'ensemble des résultats et d'un couple ordonné dont les termes appartiennent à l'ensemble des éléments sur lesquels on opère ?

...au profit des propriétés structurales

C'est justement cet éloignement de la quantité au profit des structures et des relations d'ordre qui explique le succès de la mathématique moderne dans les sciences humaines. Elle constitue pour celles-ci une typologie de structures susceptibles d'éclairer les rapports entre éléments au sein des ensembles sociaux. A la suite des travaux de Kurt Lewin ◆ et de Jacob Moreno ◆ sur les interactions sociales qui structurent la vie d'un groupe, les notions de la mathématique moderne ont été de plus en plus utilisées en psychosociologie. Les graphes, notamment, mettent en évidence les relations entre individus ou entre sous-groupes. Ainsi, même les sympathies ou antipathies liant les hommes au sein d'une société, qui paraissaient échapper, comme les manifestations essentiellement qualitatives de l'affectivité, à l'emprise des mathématiques quantitatives, sont susceptibles à l'heure actuelle d'une approche mathématique par la grâce des mathématiques modernes.

◆ K. Lewin (1890-1947) : psychosociologue américain d'origine allemande. Il a fondé la dynamique des groupes sur la topologie mathématique.

◆ J. Moreno (né en 1892) : psychosociologue américain d'origine roumaine, qui a créé le psychodrame et la sociométrie.

K. Lewin proposait, pour l'analyse des relations affectives et sociales, la notion de « champ total » englobant le sujet humain avec ses tendances et les sollicitations de l'environnement. S'inspirant déjà de la topologie, K. Lewin analysait ce « champ total » en termes de voisinages, d'enveloppements, d'intersections, de séparations, de frontières comprenant notamment les « barrières psychiques » ou inhibitions et interdictions de toutes sortes.

Puis Bavelas ◆ décrivit ces totalités au moyen de graphiques, qui sont des cas particuliers de la théorie des graphes, et mit en évidence des structures de réseaux. Dans le prolongement de ces recherches, l'interprétation mathématique de la « dynamique des petits groupes » se développe avec succès ; et c'est bénéfique en retour pour la théorie des graphes, dont certains chapitres actuels viennent en droite ligne des études sur les réseaux de communication ou de l'analyse sociométrique.

◆ Bavelas : psychosociologue américain contemporain.

Ethnologie..., topologie..., calcul matriciel, etc.

Dans une autre perspective, celle de l'anthropologie culturelle, qui étudie les sociétés « élémentaires » au sein desquelles les processus psychologiques sont indissociables des structures linguistiques, économiques, juridiques, C. Lévi-Strauss ◆ est parvenu à retrouver dans les diverses organisations de la parenté des structures algébriques de réseaux et de groupes de transformation, qu'il a pu dégager avec l'aide de mathématiciens tels A. Weil et G.-Th. Guilbaud. Bien plus, les mythes eux-mêmes recueillis par l'ethnologie font l'objet, dans le cadre de l'anthropologie culturelle, d'une étude structurale dont la mathématique moderne et les techniques de l'informatique fournissent les instruments : C. Lévi-Strauss recourant à l'analyse combinatoire dégage la logique qui commande l'organisation interne d'un mythe et ses diverses variantes générales, même quand elles sont cachées par des sédiments culturels particuliers. Ainsi, on peut analyser, à l'aide du calcul matriciel, les variantes d'un mythe : il y a une série significative d'éléments qui changent d'une manière ordonnée d'un mythe à l'autre. Dès lors, des mythes en apparence très éloignés les uns des autres se révèlent unis en profondeur par des structures identiques dont ils procèdent.

◆ Claude Lévi-Strauss : ethnologue français, né en 1908. Professeur au Collège de France, il est l'un des principaux représentants du structuralisme.

Ces directions de recherche en sciences humaines rejoignent celles de la linguistique structurale. De Saussure a fondé celle-ci lorsqu'il a défini le langage comme « système » de signes à base de distinctions et d'oppositions interdépendantes : le rapport linguistique fondamental étant une correspondance entre le signe et le sens, l'ensemble des significations d'une langue forme un système de distinctions et d'oppositions, qui en permet la compréhension.

Dès lors, la linguistique, dans la mesure où elle assimile le langage à un code, devait trouver une aide précieuse dans les techniques de l'informatique et, par suite, dans la mathématique moderne. Notamment, quand on s'est intéressé, à la suite de N. Chomsky ◆, à la

◆ Noam Chomsky : américain, né en 1928, chef de file d'une école linguistique qui s'efforce de lier par une « grammaire générative » les propriétés formelles du langage à la créativité de l'esprit humain.

génération des structures syntaxiques, on a tenté de trouver une formalisation logico-mathématique des transformations linguistiques qui possèdent un pouvoir régulateur.

« *Ensembles* » *et biologie*

Ces mathématiques conçues pour la linguistique ont des applications en dehors des sciences humaines, partout où l'on peut interpréter les phénomènes en termes de codes et de signaux. C'est le cas en particulier de la biologie, depuis la mise en évidence du « code génétique ».

La biologie utilise maintenant ces nouveaux moyens d'investigation issus de la mathématique moderne, après s'être cantonnée longtemps dans l'exploitation des méthodes statistiques. Jusqu'à ces dernières années, en effet, une des principales applications des mathématiques à la biologie, en dehors des relations fonctionnelles propres à la physiologie ou à la biochimie, était la « biométrie », autrement dit les recherches statistiques regroupant des ensembles de mesures faites sur des « populations » d'êtres vivants.

Inaugurées par les études du mathématicien belge Quételet sur la taille des conscrits, elles ont été développées et étendues à diverses espèces animales par Galton, Pearson et Volterra ◆.

◆ Galton (1822-1911) : physiologiste anglais ; Pearson (1857-1936) : mathématicien et statiticien anglais, connu par ses applications des mathématiques en biologie ; Volterra (1860-1940) : mathématicien italien qui appliqua en biologie l'analyse fonctionnelle.

Récemment encore, la biologie reposait sur les statistiques

La biométrie, pour définir rigoureusement les espèces, leurs variétés et leur filiation, les caractérise par une distribution de fréquences statistiques. Elle substitue aux relations qualitatives de ressemblances et de différences, mises en évidence par les classifications de l'histoire naturelle, un système de corrélations métriques.

La génétique mendélienne est également d'essence statistique : les caractères héréditaires, que Morgan localisa dans les gènes, ne se répartissent-ils pas comme les pièces de monnaie au jeu de pile ou face? Dans cette perspective, on peut parler d'une « loterie de l'hérédité ».

La génétique moléculaire relève de plus en plus de la mathématique moderne

Mais, si l'hérédité est essentiellement un message chimique écrit avec un alphabet moléculaire, les opérations de codage et décodage reposant sur la combinatoire des ensembles deviennent fondamentales. Ses processus n'apparaissent-ils pas, dans cette perspective, structuralement semblables à la traduction d'un message d'une langue dans une autre par l'intermédiaire d'un code, où les signes, comme dans tout langage, ne prennent un sens que par leur combinaison? Sous cet angle, la combinatoire des ensembles est inscrite au cœur même de la vie. La mathématique moderne éclaire ainsi la « logique du vivant ».

On se rend compte de l'ampleur des applications des mathématiques aujourd'hui. Elles jouent un rôle qui tend à devenir prépondérant dans tous les domaines de la recherche, qu'il s'agisse de sciences physiques, de biologie, de technologie, d'économie politique ou de sciences humaines. En ce sens, elles conditionnent désormais directement l'avenir d'une civilisation industrielle, à tel point qu'on a pu dire que, pour les nations modernes, le mot d'ordre est devenu « échec ou maths ».

« MATHEMATIQUES PURES » ET « MATHEMATIQUES APPLIQUEES »

Ces applications des mathématiques résultent-elles simplement de l'utilisation des résultats acquis par les mathématiciens et employés tels quels pour aborder les problèmes pratiques? Autrement dit, les « mathématiques appliquées » sont-elles des applications directes des « mathématiques pures »?

L'application des mathématiques pures à des phénomènes concrets n'est, le plus souvent, ni immédiate ni directe

On conçoit habituellement les « mathématiques appliquées » comme l'« application » des mathématiques pures à des phénomènes concrets, dont on veut donner une représentation rationnelle et une expression numérique calculable. Mais il faut bien en convenir : le plus souvent, cette application n'est ni immédiate ni directe. Elle exige de l'ingéniosité non seulement pour adapter les méthodes générales aux cas particuliers, mais aussi pour inventer les méthodes d'application proprement dites. Ce sont ces dernières qui constituent, en effet, l'essentiel des « mathématiques appliquées ».

Les calculs numériques

Notamment, pour effectuer un calcul numérique, le physicien, comme l'ingénieur, devra souvent employer des méthodes différentes des solutions théoriques. S'il s'agit, par exemple, de calculer une intégrale, on ne peut dans la majorité des cas concrets utiliser les méthodes classiques : la notion d'intégrale implique, en effet, l'accroissement infiniment petit de la variable, ce qui n'est pas le cas en pratique. En particulier, avec un ordinateur où ces accroissements ne peuvent être que finis et, par conséquent, discontinus, il faut avoir recours à d'autres méthodes de calcul numérique, dites « par différences finies », pour calculer une valeur approchée de l'intégrale. Le technicien pourra même préférer à une méthode rigoureuse,

parce que bien fondée théoriquement, un procédé de calcul plus rapide, surtout s'il permet d'obtenir des résultats significatifs.
De plus, les fonctions dont disposent habituellement physiciens, ingénieurs et économistes sont empiriques : elles sont données soit par l'enregistrement d'une courbe continue, soit par un tableau de valeurs discrètes, où la variable n'est pas forcément en progression arithmétique. La détermination approchée d'une intégrale revient, dans le premier cas, à l'évaluation de l'aire comprise entre la courbe enregistrée, l'axe des x et les verticales passant par les points extrêmes. Dans le cas où seuls sont connus certains points de la courbe représentant le phénomène étudié, pour calculer l'intégrale définie correspondante, on a recours, par exemple, à la « méthode des trapèzes », où l'on remplace la courbe entre deux points par un segment de droite. Quand il faut différentier, comme il y a souvent dans l'enregistrement d'un phénomène des variations irrégulières dues à l'appareillage ou à des phénomènes secondaires (des vibrations, par exemple), on est bien obligé de substituer à la notion de « dérivée », qui perd ici toute signification, celle du « taux de variation », définie par référence à une courbe stylisée obtenue à partir de la courbe expérimentale en négligeant ses irrégularités.

Les modèles mathématiques

Ainsi, dans beaucoup de cas, l'expression mathématique qui exprime un phénomène ne permet pas des calculs numériques aisés. Or à quoi sert de donner une traduction symbolique d'une expérience ou du fonctionnement d'un moteur si l'on ne peut l'utiliser, parce qu'on ne sait pas effectuer les calculs correspondants?

Pour établir un modèle pratique, il faut parfois séparer des propriétés simultanées

C'est pourquoi la tâche des mathématiques appliquées est d'élaborer non seulement des méthodes de calcul, mais aussi et surtout des instruments de traitement mathématique des phénomènes permettant d'aboutir à des expressions calculables. Notamment, il est rare que les montages techniques ou les variations économiques s'expriment sous forme d'équations différentielles simples : le mathématicien spécialisé dans la recherche appliquée tente alors d'élaborer un modèle représentatif suffisamment approché des phénomènes étudiés et exprimable par des équations plus faciles à résoudre que celles d'un système plus satisfaisant peut-être au point de vue théorique, mais difficilement utilisable par les praticiens. Pour cela, il essaye de concevoir une modification dans la structure du phénomène considéré, qui laisse intactes ses caractéristiques principales et rende possible une transcription mathématique plus simple. Une telle représentation, tout en n'étant qu'approchée, exprime les propriétés essentielles qu'il s'agit de calculer. Par exemple, dans l'étude de la

propagation de la chaleur, on rencontre des expressions compliquées provenant de ce qu'on doit alors tenir compte de deux propriétés simultanées : la matière emmagasine de la chaleur, mais aussi elle la laisse s'écouler. Dès lors, si l'on veut un modèle de ces phénomènes permettant des calculs numériques plus simples, il faut séparer leurs propriétés simultanées. Comme il est indiqué dans un ouvrage de mathématiques appliquées ◆, on substitue à un conducteur continu, pourvu de la double qualité de récepteur de chaleur et de conducteur, un ensemble discontinu de récepteurs et de conducteurs de chaleur régulièrement espacés. On fait ainsi apparaître des expressions mathématiques qui rendent possibles les calculs numériques. Déjà Aristote n'affirmait-il pas que, sur le plan de la représentation mathématique du monde physique, « on peut arriver à » d'excellents résultats en posant comme séparé ce qui n'est pas » séparé ◆ »?

◆ F. Salles : *Utilisation en technique des équations et systèmes différentiels* (Publications scientifiques et techniques du ministère de l'Air, 1959).

◆ Voir Aristote : *Métaphysique* (M, chap. 3, 1078*a*).

On se rend compte de toute l'ingéniosité qu'exige l'« art de l'ingénieur », à ce stade d'élaboration théorique : pour imaginer un modèle mathématique représentant avec fidélité un phénomène technique, il faut être bon physicien autant que mathématicien. Il en est de même pour l'économiste qui construit un modèle permettant de simuler un phénomène économique et d'étudier son évolution dans le temps à l'aide d'un ordinateur : il faut combiner judicieusement les hypothèses afin de représenter de la manière la plus approchée la réalité qu'on veut calculer, ce qui implique une connaissance précise des processus mis en jeu.

Les algorithmes employés dans les calculs numériques ne doivent pas être trop sensibles aux erreurs initiales

Le mathématicien spécialisé dans la recherche appliquée doit tenir compte aussi de l'influence sur le résultat final des hypothèses simplificatrices qui ont été nécessaires pour construire tel modèle utilisé. De ce point de vue, il est important pour lui de distinguer, avant tout calcul, ce qui est essentiel de ce qui ne l'est pas. S'il peut sembler intéressant de négliger les termes qui paraissent petits, il ne faut pas oublier qu'ils ont parfois une influence considérable, de sorte que leur suppression changerait profondément la nature des équations exprimant le phénomène correspondant. Par exemple, dans les problèmes d'hydrodynamique de la viscosité, on ne peut négliger les forces de viscosité, malgré leur faible amplitude, sous peine de commettre de graves erreurs : elles régularisent, en effet, les vitesses en s'opposant au développement des distorsions. Donc, avant de faire une simplification en supprimant un terme dans les équations du modèle utilisé, il faut examiner soigneusement son influence réelle, et cela d'autant plus que, souvent, la seule présence du terme considéré comme négligeable est beaucoup plus importante que sa grandeur.

Le mathématicien qui est spécialisé dans la recherche appliquée

se heurte à une autre difficulté : les mesures qu'il s'agit pour lui de coordonner en un système comportent toujours une marge d'erreur. Il importe, par conséquent, que les algorithmes employés dans les calculs numériques ne soient pas trop sensibles à des erreurs initiales. Sous cet angle, l'intégration et l'usage des équations intégrales sont souvent préférables aux équations différentielles. En effet, une erreur commise dans l'évaluation d'une grandeur se répercute en croissant dans les dérivées successives de celle-ci, alors que l'intégration en diminue au contraire l'influence. La supériorité de l'intégration sur la différentiation en mathématiques appliquées provient aussi, plus fondamentalement, de ce que, dans ce domaine, on se trouve toujours en présence de situations concrètes particulières : or les caractéristiques qui définissent celles-ci sont automatiquement introduites dans le calcul intégral sous forme de « conditions initiales », alors qu'il est difficile de les faire apparaître dans une expression différentielle qui est éloignée des cas particuliers par suite de sa généralité. C'est pourquoi, pour rendre compte par exemple de l'évolution d'un système physique, il est préférable de l'exprimer sous la forme du bilan énergétique des échanges avec l'environnement du temps t_0 au temps t, ce qui permet l'utilisation avantageuse du calcul intégral : le point de vue newtonien, où l'on identifie le système physique considéré à un système de forces définies respectivement par des relations du type $f = m\gamma$, est plus classique, mais il aboutit à des équations différentielles souvent difficiles à résoudre.

Les modèles analogiques permettent souvent une résolution par des procédés purement physiques

Toujours en vue de simplifier les calculs, le mathématicien spécialisé dans les recherches appliquées a recours aussi à des modèles « analogiques » : il substitue à un phénomène, dont l'étude mathématique est délicate, un autre phénomène de nature toute différente, mais qui présente avec lui une analogie de structure et dont l'étude mathématique est plus facile. Notamment, les analogies entre les montages électriques et les systèmes mécaniques permettent, dans beaucoup de cas, de substituer à la résolution d'équations difficiles une résolution par des procédés purement physiques. Des modèles analogiques sont employés par exemple en construction aéronautique, où l'on étudie mathématiquement la structure d'une aile d'avion à l'aide d'un circuit électrique qui lui est équivalent par ses caractéristiques structurales.

Même lorsqu'on est amené ainsi à reconnaître la spécificité des « mathématiques appliquées » et leur relative autonomie en ce qui concerne l'élaboration des modèles représentatifs facilitant les calculs numériques, on a tendance à penser cependant qu'elles ne sont sur le plan théorique que des conséquences des « mathématiques pures » et qu'elles s'y intègrent pleinement, dès que le mathématicien a réussi à justifier rationnellement les constructions symboliques

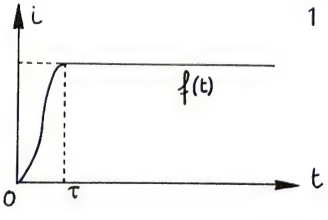
utilisées dans les sciences appliquées. Or, même dans ce cas, un hiatus doit être maintenu entre « mathématiques pures » et « mathématiques appliquées » sur le plan de la signification de leurs symboles respectifs.

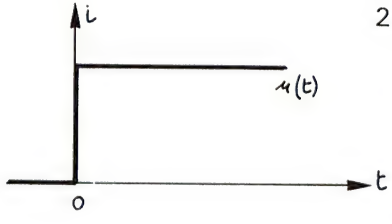
L'autonomie des « mathématiques appliquées »

Les travaux de Heaviside notamment révèlent cette double autonomie des « mathématiques appliquées » : ni leur découverte ni même leur signification ne coïncident pas toujours avec celles des structures théoriques qui les justifient. Pour faciliter les calculs des électroniciens, Heaviside a imaginé un nouveau système d'écriture des expressions différentielles, qui est devenu le « calcul symbolique » : il a fait ses preuves dans la pratique avant d'être fondé sur des bases théoriques solides. Or, quand on écrit à l'aide des symboles de Heaviside les équations différentielles qui définissent, par exemple, un servo-mécanisme ou encore le guidage d'une fusée soumise à l'action d'un faisceau radar, les formules obtenues ont une signification différente pour le pur mathématicien et pour l'ingénieur. Dans la perspective du premier, elles peuvent s'écrire algébriquement de plusieurs façons, toutes équivalentes. Il n'en est pas de même du point de vue technique, car une relation en calcul symbolique exprime un montage physique déterminé : dès lors, il ne convient pas de changer sa forme sans précautions.

La « fonction échelon unité » représente les phénomènes physiques transitoires

Heaviside a aussi imaginé la « fonction échelon unité » pour représenter les phénomènes physiques « transitoires » qui se produisent pendant un intervalle de temps très bref au cours du déclenchement d'autres phénomènes : c'est le cas, par exemple, du passage d'un courant électrique dans un circuit au moment de la manœuvre du commutateur, quand l'intensité passe brusquement de la valeur 0 à la valeur i, selon une courbe en S qui n'est jamais la même, car l'état transitoire se modifie à chaque nouveau déclenchement puisque les étincelles de contact ou de rupture transforment les caractéristiques des plots du commutateur. La « fonction échelon unité » est une idéalisation mathématique commode de la fonction correspondant à la courbe en S, essentiellement variable et insaisissable :





1. Courbe représentant la fonction qui exprime un phénomène « transitoire ».

2. Fonction « échelon unité » $\mu(t)$ *(elle est la limite de f (t)* quand τ tend vers zéro).

Elle est la limite de la fonction représentée par cette courbe en S quand la période transitoire tend vers 0. Son nom vient précisément de ce qu'elle passe brusquement du niveau 0 au niveau + 1. Puisque cette fonction est nulle pour $t < 0$ et égale à 1 pour $t > 0$, cette discontinuité caractéristique pose un problème pour le mathématicien : quelle est sa valeur pour $t = 0$? Est-ce 0, + 1 ou une valeur intermédiaire? Or habituellement, en mathématiques pures, on pense les discontinuités comme des limites de fonctions continues. Par suite, on a imaginé une fonction continue permettant par un passage à la limite de retrouver la « fonction échelon unité ». Selon cette interprétation, la « fonction échelon unité » a la valeur $\frac{1}{2}$ pour $t = 0$ ◆. Or cela n'a aucune signification pour le physicien ou l'ingénieur : pour eux, en effet, l'action commence à l'instant zéro, et à cet instant l'effet de l'action est nul ; il ne peut donc avoir la valeur $\frac{1}{2}$. Ainsi une fonction, d'abord imaginée par les physiciens pour représenter des états transitoires observés au laboratoire puis idéalisée par Heaviside pour les besoins des calculs de l'ingénieur en la « fonction échelon unité », est transformée par le mathématicien qui interprète sa discontinuité en un pur être de raison, vidé en définitive de sa signification physique initiale. Ce hiatus se retrouve dans l'étude des dérivées de la « fonction échelon unité », dites « fonctions de Dirac » et appelées aussi « fonctions percussionnelles », parce qu'elles expriment un brusque changement d'état en modifiant les conditions initiales d'un phénomène physique. Or elles sont très utilisées en électronique, comme en mécanique, où l'ingénieur ne doit jamais oublier leur signification concrète aux diverses phases des calculs, sous peine d'erreurs grossières.

◆ Voir Potier et Laplume : *le Calcul symbolique* (Paris, Hermann, 1961).

On se rend bien compte ici que des visées divergentes différencient fondamentalement les « mathématiques appliquées » et les « mathématiques pures ». Le pur mathématicien s'intéresse aux structures idéales, à leurs transformations et combinaisons virtuelles, qui suggèrent des développements théoriques pris comme fin en soi. Le spécialiste des mathématiques appliquées tente de discerner les adaptations possibles des notions théoriques aux problèmes concrets et les aptitudes des phénomènes à une stylisation permettant l'utilisation de théories mathématiques. Dans ces conditions, c'est la fidélité à l'expérience qui oriente les « mathématiques appliquées », tandis que le souci de rigueur anime les « mathématiques pures ».

Dans la recherche mathématique appliquée, l'expérience l'emporte sur la logique formelle

Ces visées respectives de la recherche pure en mathématiques et des recherches appliquées non seulement délimitent deux domaines mathématiques différents, mais encore définissent deux types

d'« objets » en mathématiques. On s'en rend compte quand on compare les interprétations de la « fonction échelon unité » et des « fonctions de Dirac », par les « mathématiques pures » et les « mathématiques appliquées ». Plus que d'« interprétations » différentes d'une même structure rationnelle, il s'agit chaque fois en réalité de deux « objets » différents. L'un a des origines physiques : il est la stylisation d'un phénomène et il n'a une signification que par référence à ce dernier. L'autre est une essence pure considérée en soi, sans être référée à des origines concrètes ou à une situation expérimentale. Même quand les « objets » respectifs des « mathématiques pures » et des « mathématiques appliquées » sont identiques au point de vue formel, c'est-à-dire quand ils ont les mêmes caractéristiques et sont rigoureusement superposables par leur structure rationnelle, ils diffèrent encore par leurs significations : l'objet des « mathématiques appliquées » est inséparable de son domaine d'application ; on y distingue notamment les quantités « négligeables » et les termes « significatifs », on tient compte des « valeurs initiales » concrètes...

Si, dans la perspective de la récente « théorie des distributions » de Laurent Schwartz, la « fonction échelon unité » et celles de Dirac apparaissent situées dans un ensemble structural qui les éclaire, l'ingénieur qui les utilise ne peut faire abstraction de leurs références physiques.

Ainsi, le domaine de la recherche mathématique appliquée est celui où l'expérience l'emporte en définitive sur la logique formelle. Les relations mathématiques y ont une signification concrète, qu'elle soit physique ou économique, voire humaine, qu'on ne doit jamais oublier sous peine d'aboutir à des échecs sur le plan pratique. Alors que les notions des « mathématiques pures » ont une signification toute rationnelle, celles des « mathématiques appliquées » se réfèrent toujours à des situations ou à des opérations effectuables sur ces dernières. Sous cet angle, leur signification est surtout technico-rationnelle.

LES MATHEMATIQUES ET LE REEL

Ce qui précède montre l'ampleur et la variété des applications des mathématiques, mais également leurs difficultés. Quelles sont les raisons de ces succès, mais aussi des résistances que semble opposer le réel à une totale mathématisation ? « Comment se fait-il, demandait » Heaviside, qu'un être physique représenté par un vecteur n'ait » pas toutes les propriétés des vecteurs ? »

Telles sont les principales questions auxquelles doit répondre une

philosophie des mathématiques. Cette correspondance entre structures mathématiques et réalité, qui à la fois s'étend sans cesse et n'est toujours qu'approchée, doit résulter de la nature profonde des mathématiques et des phénomènes auxquels on les applique.

Traditionnellement, rationalistes et empiristes s'affrontent

Le rationalisme, qui en fait un produit de l'activité rationnelle en soulignant le caractère idéal des notions mathématiques et leur nécessité, paraît triompher aisément de l'empirisme, qui y voit de simples abstractions extraites directement du réel. Mais il se heurte à de grandes difficultés dès qu'il cesse de considérer les mathématiques en elles-mêmes pour tenter d'expliquer comment ces purs produits de l'esprit s'appliquent au réel.

Dans la perspective empiriste, au contraire, cet échange de services entre les mathématiques et la physique se conçoit aisément, puisque les mathématiques ne font que rendre au monde ce qu'elles lui ont emprunté. Mais, alors, on met entre parenthèses les caractéristiques des notions mathématiques idéales et leur enchaînement nécessaire. Si l'on admet que les notions géométriques aient pu être suggérées par certaines données de la perception, comment les mathématiques pures pourraient-elles être extraites directement de l'expérience sans faire intervenir l'activité de l'esprit? Mais, si on en fait un produit de la pensée, comment la réalité matérielle pourrait-elle avoir des affinités avec des structures purement rationnelles?

Selon Dedekind ◆, « le nombre est une émanation immédiate des lois » pures de la pensée ». Mais, alors, comment peut-il écrire que les nombres « servent de moyens pour saisir plus aisément et avec plus » de précision la diversité des choses »? Dedekind se contente de juxtaposer ces deux affirmations sans expliquer leur liaison effective. Il laisse subsister entier le problème qu'Einstein avait le courage de poser : « Comment est-il possible que la mathématique, qui est un » produit de la pensée humaine et indépendante de toute expérience, » puisse s'adapter d'une si admirable manière aux objets de la réalité? » La raison humaine serait-elle donc capable, sans avoir recours à » l'expérience, de découvrir par la pensée seule les propriétés des » objets réels? »

◆ R. Dedekind (1831-1916) : mathématicien allemand dont l'activité porta essentiellement sur les nombres. Il est le premier mathématicien à avoir analysé avec rigueur le concept de nombre irrationnel et on lui doit, en particulier, la notion de « coupure ».

Il ne peut y avoir de moyen terme entre un pur produit de la pensée et le monde, à moins d'admettre une sorte d'harmonie entre eux fondée dans l'absolu. Mais, alors, on ne voit pas comment cette affinité profonde des mathématiques et du réel ne permettrait qu'une application approchée des premières au second, si étendue soit-elle. Une conception adéquate des mathématiques doit rendre compte non seulement de la concordance entre cette science et des secteurs sans cesse plus nombreux de la réalité, mais aussi du caractère approché de la mathématisation réussie, voire de l'échec auquel ont abouti les tentatives pour donner une expression mathématique à certains phénomènes.

Il est regrettable que la plupart des philosophies rationalistes aient posé le problème des mathématiques appliquées en oubliant cette contrepartie. Mais pouvaient-elles faire autrement? Entre un pur produit de la pensée et la matière, il ne saurait y avoir qu'une hétérogénéité complète ou une harmonie préétablie totale.

Pour penser les aspects divers des rapports entre les mathématiques et les sciences physiques, notamment, il ne faut donc pas commencer par poser abstraitement l'un en face de l'autre l'esprit et le monde. Mais, alors, comment éviter la tentation empiriste en les rapprochant? Si l'on peut donner à des secteurs très étendus du réel une expression mathématique et si celle-ci n'est toujours qu'approchée, ne serait-ce pas parce que les notions mathématiques ont été conçues à partir d'une stylisation d'opérations concrètes sur la matière? On comprendrait alors qu'elles puissent s'y appliquer, mais toujours imparfaitement, puisque toute stylisation implique idéalisation. Qui dit stylisation dit activité intellectuelle qui stylise, ce qui est à l'opposé de l'empirisme, sans être pour autant un retour pur et simple aux thèses rationalistes, car qui dit stylisation dit aussi stylisation d'un donné qui s'y prête et la motive.

Et à partir de quel donné pourrait-on obtenir, par stylisations successives, le nombre, concept opératoire par excellence, et les notions qui en sont dérivées par de nouvelles stylisations faites sur les opérations effectuées avec lui, sinon à partir d'opérations concrètes effectuées sur des objets quelconques, puisque le nombre s'applique justement à des objets quelconques? Les opérations qui permettent la superposition d'objets réels semblent pouvoir jouer ce rôle. En effet, elles sont à la base aussi bien de la mesure que de la constitution des collections.

Les mathématiques les plus abstraites ont un soubassement d'opérations concrètes

Cette conception des mathématiques, qui rejoint les perspectives de l'épistémologie génétique de J. Piaget ◆ et celles de Gonseth ◆, aide à comprendre les rapports de cette science et du réel : chaque fois qu'une stylisation des opérations faites sur un phénomène est réalisable selon ce modèle, c'est-à-dire, en définitive, lorsque le phénomène présente des aspects mesurables, un lien s'établit naturellement entre le réel et les mathématiques.

Les bases de l'arithmétique ont leur origine dans des opérations manuelles. C'est indéniable. L'addition et la multiplication s'y réfèrent directement : ne désignent-elles pas la répétition d'un acte simple de préhension, qui met en rapport la main et un objet quelconque? Non seulement la numération décimale est dans le prolongement des dix doigts sur lesquels on compte, mais encore la théorie des ensembles elle-même ne peut se passer de références corporelles, qui lui donnent un sens.

La correspondance biunivoque est une stylisation de l'échange un

◆ J. Piaget (né en 1896) : psychologue, pédagogue et logicien suisse, connu par ses travaux sur la psychologie de l'enfant et sur la formation des structures intellectuelles.

◆ Gonseth (né en 1890) : mathématicien et philosophe suisse ; a écrit *les Fondements des mathématiques* (1926).

contre un. La formation de sous-ensembles, leur réunion, leur ordination, etc., ont une signification parce que ce sont d'abord des opérations effectuées par la main. Quant à la permutation d'éléments symétriques, ne présuppose-t-elle pas, en plus de mouvements de la main, la symétrie corporelle?

Bien plus, c'est toute l'analyse qui repose, en définitive, sur l'activité du corps en prise sur l'environnement. Ses notions n'ont-elles pas été construites à partir des opérations arithmétiques qui leur ont légué leurs réminiscences corporelles? Comme le remarquait Meyerson : « Dans la formule d'Euler $e^{ix} = \cos x + i \sin x$, un » imaginaire figure comme exposant. Cela est assurément impossible » à transporter directement dans le réel. Mais un exposant est la » caractéristique d'une puissance, laquelle se ramène à une multi- » plication qui, à son tour, n'est qu'une addition abrégée, de telle » sorte que nous voici en présence de nos cailloux » (dont on se servait pour compter autrefois, dans les échanges commerciaux).

Puisque les nombres naturels entrent comme composants, ne serait-ce qu'à titre de scalaires ou de coefficients, dans les différentes formes du nombre généralisé, les références corporelles de l'arithmétique hantent les plus hautes spéculations mathématiques, en leur donnant une signification opératoire.

« Pour un être complètement immobile, il n'y aurait ni espace ni géométrie » (Poincaré)

Il n'y a pas de géométrie sans référence à l'activité de la main qui trace des figures : leur définition par le mouvement d'un point n'y renvoie-t-elle pas directement? Et comment définit-on l'égalité géométrique, sinon par la superposition, c'est-à-dire par une opération manuelle? Bien plus, l'espace orienté s'enracine dans l'expérience vécue du corps ordonnant l'environnement perçu selon l'avant et l'arrière, le haut et le bas, la droite et la gauche.

Il n'est pas jusqu'à la topologie, cette théorie des ensembles géométriques, qui implique directement l'activité corporelle : ne repose-t-elle pas sur les notions d'ordre, d'emboîtement, de frontières, etc., qui toutes renvoient à des opérations effectuées originairement par la main?

Ainsi, les mathématiques de l'ordre, comme les mathématiques de la quantité, se réfèrent à des opérations corporelles s'exerçant sur l'environnement. Dès lors, sous leur revêtement abstrait, elles sont liées encore au monde. En ce sens, Cavaillès a pu écrire : « Le sensible » n'est pas abandonné : ce n'est pas le quitter que d'agir sur lui. » Tout objet abstrait obtenu, par exemple, par thématisation est » un geste sur un geste..., sur un geste sur le sensible primitif. Le » champ thématique n'est donc pas situé hors du monde, mais est » transformation de celui-ci. ◆ »

◆ Cavaillès : *Méthode axiomatique et formalisme* (pages 178-179).

Cet enracinement des mathématiques dans le corps en prise sur l'environnement est le fondement, en définitive, des mathématiques

appliquées : il rend compte des succès, mais aussi des difficultés que rencontrent les tentatives de mathématisation du réel.
Elles réussissent dans la mesure où les phénomènes considérés, leurs relations et, surtout, les opérations qu'on peut effectuer sur eux se prêtent à une stylisation du même type que celle réalisable avec les opérations corporelles de superposition ou de mise en ordre des objets qui nous entourent. En particulier, les sciences physiques sont des stylisations de la matière. Le physicien réalise en laboratoire des phénomènes stylisés se rapprochant le plus possible de la loi idéale qui les définit, grâce à un montage technique qui écarte l'influence des phénomènes concomitants susceptibles de les perturber. De même, le chimiste opère sur des corps qu'il s'efforce de rendre pratiquement « purs », en éliminant autant qu'il le peut les autres substances qui y sont présentes.
En ce sens, les sciences de la nature sont la réalisation progressive, par des techniques appropriées, d'un projet de stylisation des phénomènes par mise en évidence de circuits de transformations et d'évolutions idéales. En elles-mêmes, ces stylisations du monde physique diffèrent certes de la stylisation mathématique : tandis que les mathématiques se développent à partir des opérations qui modifient l'ordre ou la disposition d'objets quelconques, les sciences physiques sont essentiellement une stylisation d'opérations techniques qui transforment les phénomènes.

La stylisation d'opérations concrètes rapproche le monde réel des spéculations mathématiques

Des échanges sont possibles entre ces stylisations, bien que leurs visées initiales soient différentes, justement parce qu'elles sont, les unes comme les autres, fondamentalement des stylisations d'opérations concrètes ; la symbolisation mathématique désignant les opérations les plus générales effectuables avec des objets quelconques, on conçoit qu'elle puisse être utilisée par tout autre stylisation d'opérations plus particulières, dans la mesure où l'on peut retrouver dans celles-ci les schémas plus généraux des opérations mathématiques. Notamment lorsque les notions physiques, chimiques ou biologiques sont définies à partir de mesures, elles ont une signification mathématique : celle d'opérateurs qui indiquent le sens de l'évolution d'un phénomène stylisé grâce à un appareillage approprié et suggèrent un programme virtuel d'actions techniques à réaliser sur lui.
Ainsi, la liaison entre les mathématiques et les sciences de la nature, difficilement concevable quand on considère séparément le monde, qu'étudient les secondes, et les concepts idéaux des premières, s'éclaire dès qu'on prend conscience que toutes ces sciences reposent sur des stylisations d'opérations concrètes, qui rendent possibles leurs relations. Dans cette perspective, on comprend à la fois que certaines structures mathématiques puissent anticiper le réel et que ce dernier résiste aussi à l'emprise des stylisations idéales.

Les lois de la physique expérimentale, qu'on a réussi à exprimer sous forme mathématique, peuvent être traitées directement selon les normes de cette stylisation. Mais les conséquences ainsi déduites mathématiquement n'ont une signification physique que si elles sont effectivement vérifiées dans les opérations de mesure.

On comprend également, dans cette perspective, qu'il y ait dans beaucoup de cas une étonnante harmonie entre les besoins de la physique et l'état d'avancement des recherches mathématiques, qui faisait dire à Hermite ◆ : « Je crois [...] qu'une découverte analytique » survient à un moment nécessaire pour rendre possible chaque » nouveau progrès dans l'étude des phénomènes du monde réel qui » sont accessibles au calcul. ◆ »

◆ Hermite (1822-1901) : mathématicien français connu par ses travaux d'analyse, la résolution générale de l'équation du cinquième degré et la transcendance du nombre *e*.

◆ *Correspondance d'Hermite et de Stieltjes* (I, p. 8).

Certes, cette corrélation provient d'abord de ce que les problèmes posés par la physique sont énoncés habituellement en fonction du développement atteint par les mathématiques. Surtout, dans la science contemporaine à forte charpente théorique, les faits nouveaux sont appréhendés la plupart du temps dans le prolongement d'une théorie qui suggère les expériences correspondantes. De plus, les instruments utilisés dans ces dernières sont la réalisation matérielle de théories physiques édifiées à l'aide des structures mathématiques connues.

Dès lors, l'expression symbolique des problèmes posés par les nouveaux phénomènes étudiés découle du progrès mathématique antérieur. Il est donc naturel, dans ces conditions, que souvent l'histoire de la physique ne pose que les problèmes que les mathématiciens contemporains peuvent résoudre ou, du moins, qui sont tels que les méthodes indispensables pour leur résolution sont appelées en quelque sorte par certains développements mathématiques déjà esquissés.

« Le démon empirique et le génie rationaliste sont à égalité de finesse » (Bachelard)

Mais cet optimisme a un fondement plus profond : il repose, en définitive, sur un lien essentiel entre les sciences de la nature et les mathématiques, qui provient de ce que les unes comme les autres se sont développées à l'ombre du corps de l'homme en prise sur son environnement. En effet, le physicien repère des structures matérielles susceptibles d'agir les unes sur les autres : ceci ne revient-il pas, en dernière analyse, à penser tout phénomène par référence aux modes d'action du corps sur les objets qui l'entourent? Ainsi, les actions physiques, en tant qu'elles sont en ce sens des homologues ou, du moins, comme un écho lointain des actions corporelles, doivent normalement pouvoir être exprimées par le symbolisme mathématique, puisque ce dernier est fondamentalement une stylisation de la coordination des opérations corporelles de préhension d'un objet quelconque. L'insertion du corps parmi les contacts matériels fonde l'assurance qu'il est toujours possible de faire apparaître, dans un

phénomène, des structures agissant les unes sur les autres un peu à la manière du corps sur son environnement et susceptibles, par conséquent, d'être stylisées selon un modèle mathématique. C'est sur cet ancrage du corps dans le monde que repose finalement la certitude à priori, qui est à la base de la physique, d'une correspondance entre les phénomènes matériels étudiés et une représentation mathématique ◆.

◆ Voir R. Boirel : *Théorie générale de l'invention* (Paris, P.U.F., 1961, pages 317 et suiv.).

LES APPLICATIONS PEDAGOGIQUES DES MATHEMATIQUES

Si les mathématiques sont une pyramide de stylisations effectuées à partir d'opérations corporelles sur l'environnement, faire des mathématiques, c'est apprendre à styliser les situations concrètes et les opérations faites sur elles. En ce sens, les mathématiques, qui sont un auxiliaire indispensable au physicien, à l'ingénieur comme à l'homme d'action contemporain, constituent également une discipline de l'esprit irremplaçable.

Les mathématiques, école de rigueur

C'est un lieu commun. Les professeurs y voient le fondement de leur vocation : l'enseignement des mathématiques a une signification pour eux dans la mesure où ils contribuent à l'épanouissement intellectuel des élèves. Ils « forment » des intelligences au sens propre du terme : ils leur donnent une structure, alors qu'elles n'étaient que virtualités, en leur apprenant à penser rigoureusement et à utiliser leurs capacités.

Le professeur de mathématiques enseigne non seulement des résultats et une technique opératoire, mais aussi, plus profondément, une méthode, la pratique de la déduction et des combinaisons rigoureuses, avec l'art de résoudre les problèmes. C'est pourquoi sa vocation l'oriente, en définitive, vers la transformation de son cours en un véritable « Discours de la méthode ».

D'abord apprendre à résoudre un problème, mais aussi à bien le poser

Tel est certainement le grand bénéfice d'une formation mathématique : celui qui a été entraîné à cette discipline de l'esprit a pris l'habitude de substituer aux difficultés confuses, rencontrées dans la vie pratique, des problèmes définis clairement.

La transformation d'une « difficulté » en « problème » la modifie profondément et, surtout, restructure la manière dont l'homme réagit en sa présence. Il ne la voit plus uniquement comme une gêne paralysante : grâce à cette vision claire de la « situation obstacle » devenue « problème à résoudre », il l'envisage d'une façon plus

objective et lui donne son importance réelle : c'est à cette condition qu'il peut adopter une attitude positive indispensable à la recherche d'une solution. Visant celle-ci, il essaye de combiner des moyens en vue de la dégager ; autrement dit, il n'est plus entièrement dominé par son embarras. C'est justement lorsqu'un handicap n'est pas envisagé comme « problème » qu'il peut y avoir conduite d'échec et, à la limite, névrose : l'individu est paralysé en présence d'une difficulté qu'il n'a pas réussi à penser sous la forme d'un problème à résoudre méthodiquement. En ce sens, on peut dire que le salut de l'homme est dans la conversion des « soucis » en « problèmes ».

Dès lors, les mathématiques, école par excellence pour apprendre à résoudre et, d'abord, à poser clairement les problèmes, peuvent contribuer efficacement à l'équilibre psychologique de l'homme d'action devant les difficultés, en disciplinant son intelligence. Ayant pris l'habitude de dégager dans une situation l'essentiel de l'accessoire, il disperse moins ses efforts. Sachant par expérience combien est importante la quatrième règle de Descartes : « Faire partout des dénombrements si entiers et des revues si générales que je fusse » assuré de ne rien omettre », il a le souci, quand il pose un problème, de n'oublier aucune donnée essentielle.
Or il se peut qu'un facteur primordial entrant dans la définition d'une situation concrète ne soit pas quantifiable. Dès lors, cette attitude de recherche prudente de toutes les données, acquise au cours des études mathématiques, conduit à ne pas s'enfermer dans le cadre étroit des seuls éléments chiffrés : reconnaître tous les aspects majeurs d'une situation, qu'ils soient quantifiables ou non, c'est cela la rigueur authentique, dont la rigueur mathématique n'est qu'un cas particulier ; elle relève autant de l'esprit de finesse que de l'esprit de géométrie.

Sur la lancée des habitudes prises au cours des études mathématiques, un esprit ainsi formé fait donc l'inventaire systématique des données : il interroge leurs propriétés opératoires, qui peuvent être utilisées comme moyens pour parvenir à une solution. Il examine aussi les diverses manières de procéder pour la construire. Au lieu de se figer dans une direction unique de recherche, il choisit, parmi d'autres possibles, la voie de solution qui lui paraît la plus efficace dans le cas présent, compte tenu de l'ensemble des données.

LES MATHÉMATIQUES MODERNES

L'initiation à « la mathématique moderne » dès l'école primaire habitue les enfants à utiliser des graphiques pour représenter leurs manipulations sur les ensembles : en particulier, les « intersections » et les « réunions » de collections deviennent apparentes par ce moyen.

Ils sont ainsi entraînés très tôt à schématiser des situations concrètes. Surtout — et c'est le grand avantage de « la mathématique moderne » — les enfants prennent ainsi conscience très vite des relations qui structurent les objets mathématiques. Dès lors, les situations concrètes peuvent être abordées, non plus d'une manière statique, mais dans la perspective de leurs transformations possibles et de leurs correspondances avec d'autres situations. N'est-ce pas une condition pour stimuler l'imagination? En ce sens, l'initiation à « la mathématique moderne » peut avoir des conséquences bénéfiques même dans les disciplines littéraires, notamment dans les compositions françaises où il s'agit d'imaginer des situations et leurs développements possibles.

L'élève, dont l'attention est orientée très tôt vers les relations entre éléments d'ensembles, sera entraîné à envisager les données d'un problème sous l'angle de leurs transformations possibles, de leurs variantes et de leurs combinaisons virtuelles, ce qui favorise la découverte des solutions construites à partir de ces données. On peut espérer que des enfants qui auront reçu un tel enseignement mathématique seront plus actifs en présence d'un problème : au lieu de rester hypnotisés par sa présentation initiale, d'emblée ils chercheront à modifier celle-ci, car ils y verront un faisceau de possibilités opératoires. Il est évident qu'une telle attitude, étendue en dehors du cadre des problèmes à l'ensemble des situations rencontrées, ne peut que stimuler la créativité.
Donc, « la mathématique moderne » ne développe pas uniquement l'aptitude à déduire rigoureusement : elle donne, avec le sens des relations, une orientation combinatoire à la pensée, qui est entraînée à rechercher systématiquement les multiples variantes d'une situation au lieu de rester enfermée dans une perspective unique.

Agir efficacement, comme innover vraiment, c'est être capable de réaliser ses projets

Si l'on veut former les intelligences créatrices dont les civilisations techniciennes ont besoin pour progresser et surmonter leurs difficultés, il faut donner aux élèves non seulement le sens des relations, mais aussi celui des conditions de réalisation : c'est alors, en effet, qu'ils réussiront à incarner leurs idées neuves dans des innovations effectives. Et cela suppose une intelligence qui a l'habitude de dégager les facteurs qui permettent d'aboutir. Or, de ce point de vue, rien ne vaut peut-être la pratique des constructions géométriques.

LA GÉOMÉTRIE

Si la géométrie euclidienne ne présente plus d'intérêt pour la recherche et s'il convient d'éliminer des programmes certains théorèmes sur les droites remarquables du triangle, auxquels on a sans doute donné

trop d'importance dans le passé, il ne faut pas oublier cependant qu'elle a été la grande éducatrice de l'esprit humain : ne lui a-t-elle pas révélé la méthode pour dégager les conditions de réalisation, appelée justement « analyse », car, comme l'indique l'étymologie de ce mot, on « remonte » alors d'une figure supposée construite à ses conditions de construction, et plus généralement d'un projet à ses conditions de réalisation?
En ce sens, les problèmes de géométrie élémentaire peuvent constituer, encore à l'heure actuelle, un exercice formateur irremplaçable, par la pratique de la méthode d'analyse qu'implique leur résolution. L'art de la démonstration consiste ici à choisir dans une figure complexe une figure partielle qui prend une signification particulière en fonction du problème à résoudre : il s'agit de discerner, parmi l'éventail des propriétés données, celle qui permet d'amorcer une suite de constructions acheminant vers une solution. Nous sommes ici très près des choix qu'assume le chercheur ou l'ingénieur, quand il opte pour un itinéraire d'investigation, mais aussi l'homme d'action, quand il adopte une tactique : c'est l'imagination des possibilités opératoires et le sens des conditions de réalisation qui les orientent.

Les futurs techniciens, comme tous ceux qui seront des réalisateurs, doivent être entraînés à penser ainsi. Pour réussir, il ne leur suffit pas de savoir déduire rigoureusement ou de combiner des relations : il leur faut aussi discerner les conditions de réalisation, ce qui implique un développement de l'imagination opératoire.
La géométrie paraît difficile aux débutants, en raison justement du rôle prépondérant qu'y joue l'imagination opératoire. Pour progresser dans une démonstration, pour résoudre un problème, ne faut-il pas avoir l'idée de tracer telle droite, l'idée de considérer telle figure? Sans intuition, on ne réussit pas dans ce domaine.

Pour développer celle-ci et permettre aux élèves de profiter pleinement des constructions géométriques, il faudrait enseigner directement les multiples manières de procéder pour effectuer telle opération géométrique fondamentale, par exemple les méthodes principales pour établir l'égalité de deux segments : pour y parvenir, on peut soit les superposer par une transformation appropriée, soit prouver qu'ils sont les côtés d'un triangle isocèle ou les tangentes à un cercle issues d'un même point ou, s'ils sont dans le prolongement l'un de l'autre, que la perpendiculaire à ces segments élevée par leur extrémité commune est bissectrice d'un angle dont les côtés passent par les autres extrémités, ou encore qu'ils sont soit des droites homologues de figures égales, soit les côtés opposés d'un parallélogramme, à moins qu'ils ne soient la somme ou la différence de segments égaux...
Il faudrait aussi enseigner directement les multiples constructions que rendent possibles les structures géométriques de base. Par exemple,

les débutants devraient savoir très vite qu'en présence d'une droite et d'un point on peut considérer aussi :
— la parallèle menée de ce point à cette droite ;
— la perpendiculaire abaissée du point sur la droite ;
— la projection du point sur la droite ;
— le symétrique du point par rapport à la droite ;
— la symétrique de la droite par rapport au point ;
— un cercle passant par le point et tangent à la droite ;
— le cercle ayant pour centre le point et tangent à la droite...
C'est en révélant ces possibilités multiples de déploiement des structures et opérations géométriques qu'on peut former rapidement des élèves capables d'être actifs dans les constructions et les problèmes de géométrie : ils y prendront des initiatives, ce qui leur permettra de s'y intéresser vraiment et de profiter pleinement de cet enseignement, qui les entraîne à pratiquer la méthode d'analyse et, en définitive, à penser leurs projets en termes de conditions de réalisation. En ce sens, « apprendre les mathématiques, ce serait en » réalité les inventer, c'est-à-dire se replacer au point d'où l'on sent » que l'on manœuvre [...] avec les *restrictions* et les *libertés* qui défi- » nissent l'action mathématique. ◆ »

◆ P. Valéry : *Cahiers* (t. XV, éditions du C.N.R.S., p. 148).

Dans cette perspective pédagogique, on verrait disparaître les démonstrations reposant sur des séries d'astuces, qu'on admire sans doute mais qui découragent les élèves quand elles ne sont pas justifiées logiquement. Beaucoup se souviennent, même cinquante ans après comme Meyerson ◆, de la « difficulté avec laquelle ils retrouvaient » les droites à tracer, difficulté qui n'était [...] que la traduction de » ce que les figures avaient d'inattendu ». La démonstration du théorème de Pythagore n'aurait pas été le fameux « pont aux ânes » si, au lieu du début traditionnel : « Abaissons la hauteur sur l'hypoténuse et prolongeons-la », on avait commencé en indiquant les raisons de cette construction. Alors, les élèves n'y auraient pas vu un tour de prestidigitation intellectuelle : ils se seraient rendu compte qu'ils auraient pu en avoir l'idée, à condition de raisonner et de dégager les problèmes généraux, avec leur éventail de solutions possibles, impliqués par le problème particulier considéré et dont dépend sa résolution. De quoi s'agit-il en effet, sinon de démontrer qu'une surface est égale à la somme de deux autres ? Or, pour établir une égalité de la forme $A = B + C$, on peut soit essayer de réaliser la somme $B + C$ afin de montrer ensuite qu'elle est égale à A, soit diviser A en deux parties susceptibles d'être égales respectivement à B et à C. La deuxième voie étant la plus réalisable dans le cas présent, on s'oriente vers le choix d'une droite qui divise le carré construit sur l'hypoténuse en deux rectangles pouvant être reliés aux petits carrés construits sur les autres côtés, afin d'avoir la possibilité d'établir éventuellement les égalités visées par l'énoncé. Alors,

◆ Meyerson (1859-1933) : philosophe des sciences français.

le tracé de la hauteur s'impose ou, du moins, apparaît comme une tentative suggérée par cette orientation de recherche raisonnée : en effet, ne passe-t-elle pas par le sommet de l'angle droit, point commun des deux petits carrés?
Ainsi donc, l'idée de tracer cette droite vient naturellement à l'esprit quand on replace le début de la démonstration du théorème de Pythagore dans la perspective d'un problème plus général qu'il implique, « prouver une égalité de la forme A = B + C », avec l'éventail de ses diverses solutions possibles.

L'ALGÈBRE

Cela peut s'appliquer également à l'enseignement des autres branches des mathématiques, en particulier à l'algèbre. Par exemple, pour démontrer la méthode qui permet de résoudre une équation du deuxième degré, pourquoi commence-t-on traditionnellement en déclarant d'une manière impérative : « Divisons (ou multiplions) les deux membres par le coefficient de x^2 »? Les élèves comprendraient mieux et seraient conquis par les possibilités opératoires du calcul algébrique si on profitait de l'occasion pour révéler celles-ci. De quoi s'agit-il, en effet? On cherche à résoudre une équation du deuxième degré alors qu'on ne sait encore résoudre que celles du premier degré. On va donc essayer de se ramener à la résolution d'expressions du premier degré, en transformant l'équation du deuxième degré. Pour cela, on va tenter de mettre cette dernière sous la forme du produit de deux expressions du premier degré ou sous la forme du carré d'une expression du premier degré.
La première voie ne donnant rien, on cherche donc à faire apparaître un carré parfait : c'est pour cela qu'on a l'idée soit de multiplier, soit de diviser les deux membres par le coefficient a de x^2. Dans le premier cas, on obtient le carré a^2x^2. Dans le second cas, on dégage x^2. Et l'on continue à transformer l'équation jusqu'à ce qu'apparaisse le carré d'une expression du premier degré, en exploitant notamment la relation classique $(a + b)^2 = a^2 + b^2 + 2\,ab$...

On peut d'ailleurs raisonner autrement pour cheminer méthodiquement vers la résolution d'une équation du deuxième degré. En effet, résoudre une équation, n'est-ce pas essentiellement déterminer x? Dès lors, cette résolution se rattache au problème plus général : déterminer un nombre x qui entre dans telles relations. Il s'agit donc, au fond, d'arriver à écrire $x = n$. Comme l'indique Bouasse dans son « Cours de mathématiques générales », « poser $x = n$ ne fait que » changer de notation. Je pose donc $x = n + m$, et je substitue : » $n^2 + m^2 + 2\,nm + np + mp + q = 0$. Cette équation doit être » satisfaite identiquement. J'ai deux arbitraires, n et m ; je vais » m'en servir au mieux : après avoir calculé l'une de ces quantités,

» l'autre doit être calculable par une équation du premier degré que,
» par hypothèse, je sais résoudre.
» J'écris : $(2\,n + p)m + n^2 + m^2 + np + q = 0$.
» Je pose $n = -\frac{p}{2}$, ce qui permet de faire disparaître m et de calculer
» $m^2 = \frac{p^2}{4} - q$ et $m = \pm\sqrt{\frac{p^2}{4} - q}$,
» d'où $x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$. »

Ici, on utilise les propriétés mêmes du symbolisme algébrique, qui permet, par des substitutions judicieuses, d'écrire sous différentes formes des expressions équivalentes.
L'élève auquel on dévoile ainsi le soubassement logique des constructions qui le déconcertent, en replaçant les problèmes particuliers dans la perspective des problèmes plus généraux qu'ils impliquent avec leurs diverses solutions possibles, non seulement n'est pas rebuté par cette discipline, mais encore se passionne très vite pour elle. Le grand principe pédagogique n'est-il pas, en mathématiques comme ailleurs, de faire prendre conscience aux enfants que ce qu'ils apprennent est à leur portée et qu'ils sont capables d'obtenir quelque succès dans les domaines étudiés?
En procédant ainsi, le professeur enseigne, en plus des théorèmes, la façon dont il convient de raisonner pour trouver les démonstrations elles-mêmes. Et c'est cela qui est éminemment formateur et qui donne le goût des mathématiques, en révélant aux élèves que les constructions ou combinaisons sur lesquelles elles reposent sont à leur portée.

AU-DELA DU RAISONNEMENT

Un enseignement des mathématiques axé de cette manière sur la résolution des problèmes et la découverte méthodique des constructions à effectuer pour progresser vers un résultat est plus qu'une présentation de propriétés remarquables : il constitue une véritable formation de l'esprit. Il habitue à envisager les situations sous l'angle de leurs combinatoires possibles et des conditions qu'elles offrent pour réaliser des projets.
En ce sens également, les mathématiques ne sont pas seulement une discipline de l'esprit : bien qu'abstraites, elles peuvent contribuer efficacement à former des attitudes intellectuelles bénéfiques pour l'action. Les discussions actuelles sur la pédagogie des mathématiques ne concernent pas uniquement l'enseignement de cette science et celui des autres sciences qui utilisent son langage : c'est l'avenir même de la civilisation technicienne qui est en jeu. En effet, ce n'est que si les hommes sont créatifs, autrement dit si on leur a appris à

être actifs en présence des situations, qu'ils seront pleinement des hommes à l'ère des supermachines de l'informatique.

Les applications pédagogiques des mathématiques sont peut-être moins spectaculaires que leurs applications techniques, mais c'est d'elles que dépend, en définitive, la mise au service de l'homme de leurs utilisations pratiques.

Les mathématiques ne forment pas seulement le raisonnement : elles impliquent une esthétique qui affine la sensibilité. Poincaré n'écrivait-il pas : « On peut s'étonner de voir invoquer la sensibilité à » propos de démonstrations mathématiques qui, semble-t-il, ne » peuvent intéresser que l'intelligence. Ce serait oublier le sentiment » de beauté mathématique, de l'harmonie des nombres et des formes, » de l'élégance géométrique. C'est un vrai sentiment esthétique que » tous les vrais mathématiciens connaissent. Et c'est bien là de la » sensibilité. ◆ »

◆ H. Poincaré : *Science et méthode*, p. 57.

Ce n'est pas étonnant, puisque les mathématiques sont un mode de stylisation et que l'art est un style ajouté à la nature. Cette racine commune à l'activité artistique et aux mathématiques n'a-t-elle pas été entrevue par les Grecs, par Platon, notamment, qui identifiait Beauté, Nombre et Idée?

Les propriétés esthétiques des figures géométriques et des courbes mathématiques sont indéniables. Remarquées très tôt, elles sont exploitées notamment par les mosaïques géométriques des Grecs et les polygones des Arabes, où une même ligne forme par ses entrelacements des étoiles, des carrés et jusqu'à des pentagones. Plus près de nous, des peintres « abstraits » puisent leur inspiration dans les combinaisons de courbes et de volumes pour exprimer l'aspiration à un ordre absolu. De même, c'est par des compositions géométriques que Vasarely obtient les effets optiques de l'« op'art » (optical art) qu'il a créé.

Que de courbes mathématiques sont gracieuses en effet ! La cycloïde, notamment, n'a-t-elle pas été appelée l'« Hélène de la géométrie »? Quant à la spirale logarithmique, souvent utilisée dans l'art du meuble ou en architecture, Bernoulli avait demandé qu'on la gravât sur son tombeau avec l'inscription : « Cette spirale merveilleuse me » plaît si étonnamment par ses propriétés singulières et admirables » que je puis à peine me rassasier de sa contemplation. »

Dans ces conditions, on comprend que Leibniz, reprenant la formule platonicienne de saint Augustin (« le nombre est vivant dans l'Art »), ait pu affirmer : « L'Art est la plus haute expression d'une arithmé- » tique intérieure et inconsciente. » Une intuition du même type a guidé bien des artistes, peintres ou architectes, dès l'Antiquité, dans leurs recherches de rapports qui commanderaient l'agencement des structures esthétiques. C'est notamment le cas du fameux « nombre d'or ». Cette proportion, qui caractérise les côtés d'un rectangle agréable à regarder, où la grande dimension est à la petite comme

la somme des deux est à la grande, soit $\frac{a}{b} = \frac{a+b}{a}$, était déjà connue des Égyptiens et des Grecs. Ce rapport doit son succès, semble-t-il, au fait qu'il est à l'image même de la vie : il permet d'engendrer une infinité de figures semblables les unes à partir des autres. En particulier, si l'on joint à un « rectangle d'or » un carré, le nouveau rectangle obtenu est également un « rectangle d'or », et ainsi de suite indéfiniment.

De nos jours, Le Corbusier a imaginé, à partir des propriétés du « nombre d'or », un « Modulor », c'est-à-dire un système de mesure pour la mise en proportions harmonieuses des ouvrages d'architecture. Les mesures du « Modulor » constituent une suite de Fibonacci dans laquelle chaque terme, obtenu en multipliant le précédent par le « nombre d'or », est la somme des deux termes qui le précèdent. Les architectes et les artistes du Bauhaus ◆ ont également utilisé de semblables proportions.

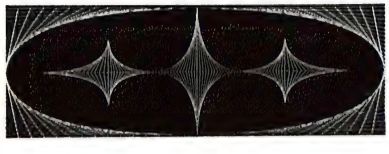
Si les arches des ponts sont gracieuses, n'est-ce pas parce que les courbes qui ont des propriétés mécaniques appréciées en résistance des matériaux sont en même temps agréables à regarder ? C'est pourquoi l'esthétique industrielle se fonde sur de semblables correspondances pour concevoir de belles structures techniques.

Quand on sait le rôle croissant que joue le « design » dans les sociétés modernes, on se rend compte que ce nouvel aspect de l'environnement technologique est également conditionné par les mathématiques. Sous cet angle, l'esthétique industrielle apparaît comme une autre application actuelle des mathématiques.

En définitive, celles-ci sont partout présentes dans les civilisations industrielles : au commencement et au cœur même de leurs progrès, car, sans mathématiques, pas de science ni de technologie avancée. Et elles contribuent à orienter leur avenir, puisque d'une formation mathématique adéquate dépend l'essor d'un véritable humanisme des techniques.

René Boirel.

◆ Bauhaus : célèbre école d'architecture allemande fondée en 1919 par Gropius. Axée sur la recherche du fonctionnel, elle a exercé une influence considérable non seulement sur la construction, mais encore sur l'esthétique industrielle et même l'art abstrait.

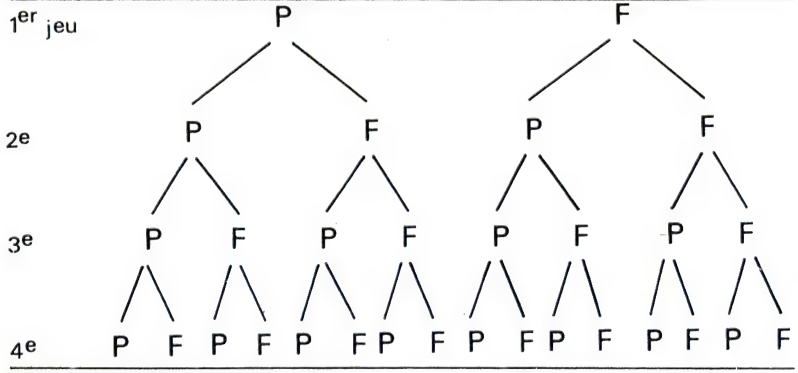


ARBRE pages 417, 425.

Schéma utilisé pour dresser la liste des parties d'un ensemble, faire un dénombrement (en calcul de probabilités, par exemple)...

Exemples

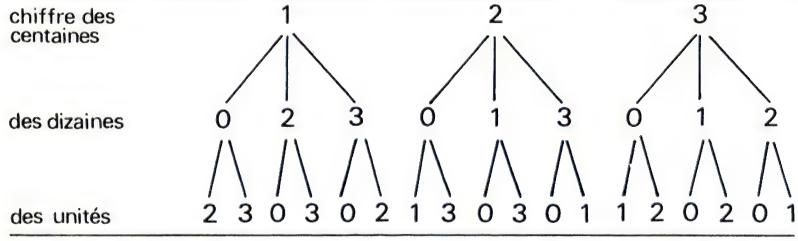
1) L'arbre donnant toutes les possibilités de jeux en jouant quatre fois consécutives à pile ou face est le suivant :



P représente *pile*,
F représente *face*.

Il y a seize « branches », chacune donnant les résultats rencontrés successivement au cours de la partie. Ce sont les seize seules possibilités. Si on cherche, par exemple, combien de fois on peut obtenir deux fois P et deux fois F, on lit les branches : PPFF, PFPF, PFFP, FPPF, FPFP, FFPP (on en trouve 6).

2) Arbre donnant les nombres de trois chiffres (les trois chiffres étant distincts) formés avec les quatre chiffres 0, 1, 2, 3 :



Il y a dix-huit nombres répondant à la question ; pour les lire, il n'y a qu'à suivre chacune des branches : 102, 103, 120, etc.

3) Arbre des parties d'un ensemble, dans le cas d'un ensemble à trois éléments $E = \{ a, b, c \}$. Pour chaque élément de E, si on se pose la question : « Appartient-il à A ? », nous n'avons que deux réponses possibles : « oui » ou « non ». On peu faire un schéma, en représentant, par exemple par — la réponse oui et par .... la réponse non.

On commence par a ; puis, sachant que a appartient à A, il y a encore 2 possibilités pour b ou, sachant que a n'appartient pas à A, il y a aussi 2 possibilités pour b...

Pour 3 éléments, on obtient ainsi 8 extrémités, chacune correspondant à un seul chemin suivi et à une seule partie de E.

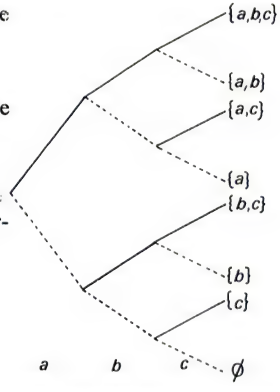
Le chemin — — — correspond à une partie A telle que a appartient à A, b aussi, c aussi, donc

$$A = \{ a, b, c \}.$$

Le chemin — — ... correspond à une partie A telle que a appartient à A, b aussi, mais pas c, donc

$$A = \{ a, b \}.$$

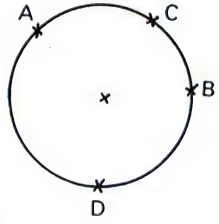
Le chemin — ... — correspond à une partie A telle que a appartient à A, b n'appartient pas à A, c appartient à A, donc $A = \{ a, c \}$.



ARC

Arc de cercle

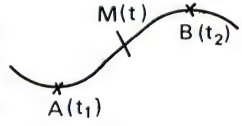
Ensemble des points d'un cercle situés entre deux points A et B de ce cercle (qui sont appelés les extrémités de l'arc). Deux points distincts A et B déterminent deux arcs de cercle : les arcs $\overset{\frown}{ACB}$ et $\overset{\frown}{ADB}$.



Arc de courbe

Étant donné une courbe plane (C) continue, dont une représentation paramétrique est $\begin{cases} x = f(t) \\ y = g(t) \end{cases}$

on appelle arc de la courbe, d'origine A et d'extrémité B, l'ensemble des points de la courbe correspondant aux valeurs de t qui sont comprises entre t_1 (valeur de t pour le point A) et t_2 (valeur de t pour le point B).



On définirait de la même façon un arc de courbe continue dans l'espace, la représentation paramétrique d'une courbe dans l'espace étant de la forme :

$$n = f(t) \qquad y = g(t) \qquad z = h(t).$$

ARCHIMÉDIEN

Soient les deux nombres réels positifs 3,5 et 17 ; il est possible de trouver un nombre entier naturel n dont le produit par 3,5 soit supérieur à 17 ; 5 répond à la question, car : $5 \times 3{,}5 = 17{,}5$.

De façon générale, quels que soient les nombres réels positifs a et x, il est toujours possible de trouver un entier naturel n tel que le produit $n \,.\, x$ soit supérieur à a :

$$\forall\, a \in \mathbb{R}^+, \forall\, x \in \mathbb{R}^+, \exists\, n \in \mathbb{N} \mid 0 \leqslant a \leqslant n \,.\, x.$$

On dit que $\mathbb{R}$ est archimédien.

$\mathbb{Z}$ et $\mathbb{Q}$ sont aussi archimédiens.

ARÊTE : voir Polyèdres

ARITHMÉTIQUE

Partie des mathématiques qui étudie l'ensemble $\mathbb{N}$ des nombres entiers naturels, c'est-à-dire positifs (0, 1, 2, 3, 4, etc.), et $\mathbb{Z}$, ensemble des entiers relatifs, c'est-à-dire positifs et négatifs ($\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots$), et les opérations dans ces ensembles, divisibilité, nombres premiers, etc.

ARRANGEMENT (Permutation) pages 285, 327.

Avec répétition ◆

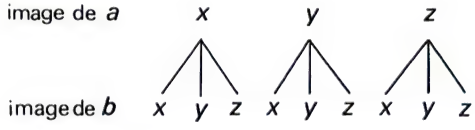
◆ Lire d'abord Application, Image et Arbre.

Cherchons à former, avec les lettres x et y, tous les mots possibles de 3 lettres, chaque lettre x ou y pouvant figurer une ou plusieurs fois ; on obtient 8 mots :

$$xxx,\ xxy,\ xyx,\ yxx,\ xyy,\ yxy,\ yyx,\ yyy$$

On dit qu'il y a 8 arrangements, avec répétition, des 2 lettres x et y, prises 3 par 3. On préfère actuellement définir ces arrangements à l'aide des applications d'un ensemble vers un autre ◆, de la façon suivante. Soit $E = \{ a, b \}$ et $F = \{ x, y, z \}$; supposons vouloir chercher toutes les applications de E dans F. Une telle application est connue dès que l'on connaît l'image de a et l'image de b. Or nous avons 3 choix possibles pour l'image de a : x, y ou z ; de même, nous avons 3 choix possibles pour l'image de b (a et b peuvent avoir la même image) ; pour chaque choix de l'image de a, nous avons 3 choix possibles de l'image de b, donc au total $3 \times 3 = 3^2$ choix possibles pour le couple (image de a, image de b), donc 9 applications de E dans F.

◆ Voir Permutation.

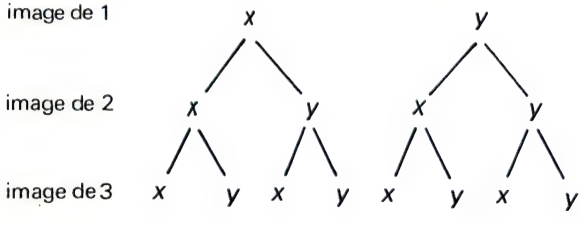


Si E est un ensemble à p éléments et si F est un ensemble à n éléments, il y a n^p applications de E dans F (puisqu'on peut choisir l'image de chaque élément de E parmi les n éléments de F).

Dans le cas où E est $\{ 1, 2, 3, \ldots, p \}$, une application de E dans F est appelée un arrangement avec répétition de n éléments pris p à p (il y en a n^p).

Par exemple, $E = \{ 1, 2, 3 \}$ $\quad F = \{ x, y \}$.

Les applications de E dans F sont au nombre de $2^3 = 8$; l'arbre permettant de les obtenir est représenté ci-contre (on reconnaît alors notre premier exemple).

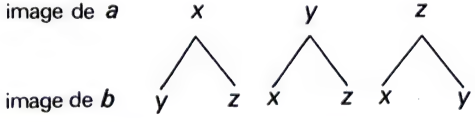


On confond encore souvent l'application f avec le résultat de l'application $f(x)$, et il arrive qu'on appelle encore arrangement avec répétition de n éléments pris p à p le résultat de l'application (c'est-à-dire les images). Par exemple, on dira que, dans l'exemple ci-dessus, il y a 8 arrangements, avec répétition, des 2 éléments x et y pris 3 par 3 et que ces arrangements sont : xxx, xxy, xyx, xyy, yxx, yxy, yyx, yyy.

Sans répétition

Si $E = \{ a, b \}$ et $F = \{ x, y, z \}$, supposons que l'on cherche maintenant le nombre d'applications injectives ◆ de E dans F. L'image de a peut être x, y ou z, celle de b doit être distincte de celle de a ; donc, si celle de a est x, celle de b ne peut plus être que y ou z ; il y a 3 choix possibles de l'image de a, mais, après chaque choix de l'image de a, il n'y a plus que 2 choix possibles de l'image de b. Il y a donc 6 choix possibles seulement pour le couple (image de a, image de b).

◆ Application telle que deux éléments distincts de E ont deux images distinctes dans F.



Si E a p éléments et F a n éléments ◆, pour qu'une application de E dans F puisse être injective, il faut pouvoir associer aux p éléments de E p images dans F qui soient toutes distinctes.
Il faut donc $p \leqslant n$.
Dans ces conditions, nous avons :

n choix possibles pour l'image de a_1 ;

$n - 1$ choix possibles pour l'image de a_2 ;

$n - (p - 1)$ choix possibles pour l'image de a_p.

◆ $E = \{ a_1, a_2, \ldots, a_p \}$
$F = \{ b_1, b_2, \ldots, b_n \}$.

Le nombre des applications injectives de E à p éléments dans F à n éléments est égal à :

$$n(n-1)(n-2) \ldots (n-p+1)$$

(produit de p entiers consécutifs décroissant à partir de n).
Si $E = \{ 1, 2, \ldots, p \}$: une telle application est appelée arrangement sans répétition de n éléments pris p à p. Le nombre de ces arrangements est noté A_n^p ◆.

◆ $A_n^p = n(n-1)(n-2) \ldots (n-p+1)$.

Par exemple :

$$E = \{ 1, 2, 3 \} \qquad F = \{ a, b, c, d, e \}$$

Le nombre des arrangements des 5 éléments de F pris 3 par 3 est :

$$A_5^3 = 5 \times 4 \times 3 = 60$$

Confondant souvent résultat d'une application et application elle-même, on dit que ces arrangements sont *abc*, *abd*, *abe*, ...

ASSOCIATIF, ASSOCIATIVITÉ ◆ pages 24 et 25, 56, 73 à 75, 86, 267, 270.

◆ Lire d'abord Composition (loi de) et l'article « l'Algèbre », page 24.

Si l'ensemble E est muni d'une loi de composition interne notée ✳, cette loi permet de composer deux éléments de E ; pour composer trois éléments de E,

on compose d'abord deux d'entre eux, puis le résultat ainsi trouvé avec le troisième. Par exemple, pour composer a, b et c (dans cet ordre), on a le choix entre composer d'abord b et c et composer a avec le résultat ainsi trouvé, ce qui s'écrit $a \divideontimes (b \divideontimes c)$, ou composer a et b et le résultat ainsi trouvé avec c, ce qui s'écrit $(a \times b) \times c$.
La loi $\divideontimes$ est dite associative si, quels que soient les éléments a, b, c de E, les deux éléments calculés ci-dessus sont égaux :

$$a \divideontimes (b \divideontimes c) = (a \divideontimes b) \divideontimes c$$

Cette propriété est vérifiée pour l'addition dans $\mathbb{Z}$ (entiers relatifs, positifs et négatifs) :

$$a + (b + c) = (a + b) + c$$

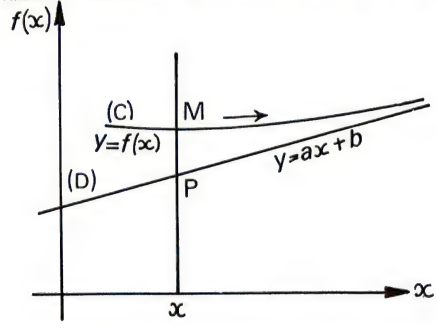
quels que soient les nombres entiers a, b et c.
Elle est également vérifiée pour la multiplication dans $\mathbb{Z}$, mais elle ne l'est pas pour la soustraction, car, par exemple :

$$\begin{cases} (3 - 2) - 1 = 0 \\ 3 - (2 - 1) = 2 \end{cases}$$

Dans le cas d'une loi associative, on note simplement $a \divideontimes b \divideontimes c$ le composé des trois éléments a, b, c.

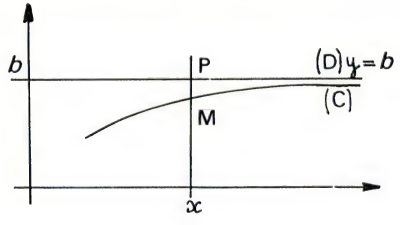
ASYMPTOTE

Si (C) est une courbe (plane) d'équation $y = f(x)$, on dit que la droite (D) d'équation $y = ax + b$ est asymptote à la courbe si la différence $f(x) - (ax + b)$ a pour limite 0 quand x tend vers $+\infty$ ou $-\infty$.



Cela revient à dire que, si l'on appelle M le point de la courbe d'abscisse x et P le point de la droite de même abscisse x, la distance de M à P a pour limite 0 quand le point M s'éloigne indéfiniment sur la courbe.

Recherche des asymptotes à une courbe ◆



1. — Si la fonction f admet une limite finie b quand x tend vers $+\infty$ ou vers $-\infty$, la droite (D) d'équation $y = b$ est asymptote à la courbe (car la différence $f(x) - b$ a pour limite 0 quand x tend vers $+\infty$ ou $-\infty$).

◆ Nous donnons les formules sans faire les démonstrations.

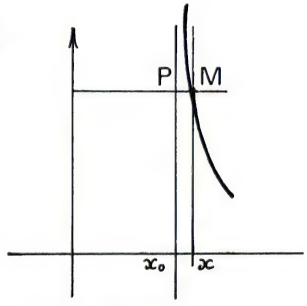
2. — Si la fonction admet une limite infinie, quand x tend vers $+\infty$ (ou $-\infty$), pour voir si sa courbe représentative admet une asymptote, on cherche la limite de $\frac{f(x)}{x}$ quand x tend vers $+\infty$ (ou $-\infty$).

— *Si cette limite est finie non nulle* et a sa valeur, on cherche alors la limite de $f(x) - ax$.

— *Si cette dernière limite est finie* et b sa valeur, il y a effectivement une asymptote d'équation $y = ax + b$ (a et b étant les 2 nombres déterminés ci-dessus). On qualifie souvent cette asymptote d'oblique, car elle n'est parallèle ni à Ox ni à Oy.

— Dans le cas où $\frac{f(x)}{x}$ admet une limite finie a, non nulle, mais où $f(x) - ax$ n'admet pas de limite finie, on dit qu'il y a une direction asymptotique (parallèle à la droite d'équation $y = ax$), mais pas d'asymptote.

Il y a un autre type d'asymptote pour une courbe (C), quand x tend vers une valeur x_0 (finie), si la fonction $f(x)$ a pour limite $+\infty$ (ou $-\infty$) ; la courbe admet pour asymptote la droite d'équation $x = x_0$.



La distance des deux points M et P, respectivement de la courbe (C) et de la droite (D), situés sur une même parallèle à Ox, a pour limite 0 quand x tend vers x_0, puisque cette distance est égale à $x - x_0$ si $x > x_0$ ou à $x_0 - x$ si $x_0 > x$.

AUTOMORPHISME : voir Homorphisme

AXE : voir Cône, Coordonnées, Cylindre, Surface de révolution, Tore

AXIOME page 60.

Au départ d'une théorie mathématique, il y a :
— d'une part, des termes primitifs que l'on ne définit pas ; par exemple, pour la géométrie : point, droite, plan, etc. ; pour les ensembles : ensembles, éléments, etc. ;
— d'autre part, des axiomes ou propositions vraies par définition.

Axiomes de Péano ◆
La définition de l'ensemble des entiers naturels $\mathbb{N}$ repose sur les :
— Termes primitifs : 0, entier naturel, successeur ;
— Axiomes :

1) 0 est un entier naturel ;

2) tout entier naturel a un successeur ;

3) deux entiers naturels ayant même successeur sont égaux ;

4) 0 n'est le successeur d'aucun entier naturel ;

5) si une partie P de $\mathbb{N}$, contenant 0, est telle que le successeur de tout élément de P appartienne à P, alors cette partie P de $\mathbb{N}$ est égale à $\mathbb{N}$.

Ces axiomes ne sont pas toujours des propriétés évidentes mais sont des énoncés posés à priori, qui doivent constituer un système cohérent (non contradictoire) à partir duquel sera édifiée la théorie.

◆ G. Péano (1858-1932) : mathématicien italien qui inventa un langage mathématique universel tout en symboles et découvrit une courbe plane dont les points ont pour coordonnées des fonctions continues d'un paramètre et qui passe par tous les points intérieurs à un carré donné.

B

BANACH : voir Norme, Normé, Suite de Cauchy

BARYCENTRE

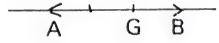
Étant donné n points de l'espace $A_1, A_2, \ldots, A_n$ et n nombres réels $\alpha_1, \alpha_2, \ldots, \alpha_n$ dont la somme n'est pas nulle, il existe un point G de l'espace et un seul tel que les vecteurs $\overrightarrow{GA_1}, \overrightarrow{GA_2}, \ldots, \overrightarrow{GA_n}$ soient liés par la relation :

$$\alpha_1\overrightarrow{GA_1} + \alpha_2\overrightarrow{GA_2} + \ldots + \alpha_n\overrightarrow{GA_n} = \vec{0}$$

Ce point G est appelé barycentre des points $A_1, A_2, \ldots, A_n$, affectés respectivement des coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$.

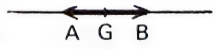
Exemple : si on se donne les deux points A et B et les deux nombres réels 1 et 2, il existe un seul point G tel que :

$$1\,\overrightarrow{GA} + 2\,\overrightarrow{GB} = \vec{0}$$



(il est sur la droite AB, puisque les deux vecteurs $\overrightarrow{GA}$ et $\overrightarrow{GB}$, vérifiant $\overrightarrow{GA} = -2\overrightarrow{GB}$, sont colinéaires ; il est situé aux deux tiers de AB à partir de A).

Si on affecte A et B des coefficients 2 et 2, le point G tel que $2\,\overrightarrow{GA} + 2\,\overrightarrow{GB} = \vec{0}$ est au milieu de AB (il en est d'ailleurs de même si on affecte A et B de coefficients égaux, pourvu qu'ils soient différents de 0).



Propriétés

Si O est un point quelconque de l'espace, on démontre que le vecteur $\overrightarrow{OG}$ est égal à :

$$\frac{\alpha_1\overrightarrow{OA_1} + \alpha_2\overrightarrow{OA_2} + \ldots + \alpha_n\overrightarrow{OA_n}}{\alpha_1 + \alpha_2 + \ldots + \alpha_n}.$$

Si O est le centre d'un repère (Ox, Oy, Oz), cette dernière égalité permet de calculer les coordonnées (x, y, z) de G si on connaît celles des points $A_1, A_2, \ldots, A_n$.

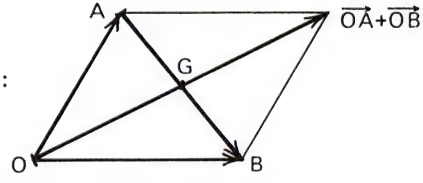
x, y, z sont donnés par les formules :

$$x = \frac{\alpha_1 x_1 + \ldots + \alpha_n x_n}{\alpha_1 + \ldots + \alpha_n} \qquad y = \frac{\alpha_1 y_1 + \ldots + \alpha_n y_n}{\alpha_1 + \ldots + \alpha_n} \qquad z = \frac{\alpha_1 z_1 + \ldots + \alpha_n z_n}{\alpha_1 + \ldots + \alpha_n}$$

si on appelle (x_1, y_1, z_1) les coordonnées de $A_1, \ldots, (x_n, y_n, z_n)$ celles de A_n.

Par exemple, si G est le barycentre des deux points A et B affectés des coefficients 1 et 1 :

$$\overrightarrow{OG} = \frac{\overrightarrow{OA} + \overrightarrow{OB}}{1+1} = \frac{\overrightarrow{OA} + \overrightarrow{OB}}{2}.$$



Les coordonnées du point G sont respectivement :

{ la demi-somme des abscisses de A et de B
{ la demi-somme des ordonnées de A et de B
{ la demi-somme des cotes de A et de B

Si les n coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ sont tous égaux, le point G est appelé équibarycentre des n points, ou encore centre de gravité. L'équibarycentre de deux points A et B est le milieu de AB ; l'équibarycentre de trois points A, B, C non alignés est le centre de gravité du triangle ABC.

Pour trouver le barycentre G de n points $A_1, A_2, \ldots, A_n$ affectés des coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$, on peut commencer par chercher le barycentre g d'un certain nombre p de ces points, par exemple $A_1, A_2, \ldots, A_p$, affecter g d'un coefficient égal à la somme des coefficients de $A_1, A_2, \ldots, A_p$ et chercher le barycentre de g et des autres points affectés de leurs coefficients.

Ainsi, pour chercher le barycentre de trois points A, B, C non alignés, affectés des coefficients 1, 1, 1, on peut :

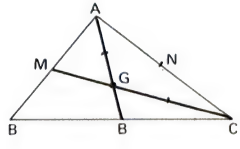
— chercher le barycentre des deux points A et B affectés des coefficients 1 et 1 (c'est donc leur milieu M) ;

— affecter le point M du coefficient 1 + 1 ;

— chercher le barycentre de C affecté du coefficient 1 et celui de M affecté du coefficient 2 ; ce point G est donc tel que :

$$1\,\overrightarrow{GC} + 2\,\overrightarrow{GM} = \vec{0}.$$

Il se trouve donc sur la médiane CM aux deux tiers à partir de C.



On aurait pu faire le même raisonnement à partir du milieu N de AC avec le point B, ou du milieu P de AB avec le point C. Le point G (barycentre de trois points affectés de deux coefficients égaux) est unique, donc les trois procédés donnent le même résultat, ce qui démontre que les trois médianes d'un triangle sont concourantes.

BASE

Base d'un espace vectoriel : voir l'article « l'Algèbre linéaire », page 65.

Base d'un logarithme : voir Logarithme.

Base d'un système de numération : voir Numération.

Base d'une topologie ◆

Les ouverts de $\mathbb{R}$ pour la topologie de l'ordre ◆ sont toutes les réunions possibles d'intervalles ouverts $]a, b[$ (a et b étant des nombres réels finis). Si on considère la famille B de tous les intervalles ouverts de $\mathbb{R}$, n'importe quel ouvert de $\mathbb{R}$ se met sous forme de la réunion d'éléments de B.

◆ Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire l'article « les Ensembles », Intervalle et Topologie.

◆ Mot du vocabulaire de topologie.

Par exemple, l'ouvert $]-2, +3[\cup]5, 6[$ est réunion des deux éléments $]-2, +3[$ et $]5, 6[$ de B. On dit que B constitue une base de la topologie de l'ordre de $\mathbb{R}$.

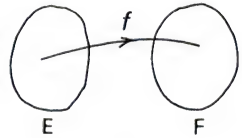
De manière générale : soit E un espace topologique ; une base de la topologie de E est un sous-ensemble B de parties de E tel que tout ouvert de E soit réunion d'éléments de B.

BERNOULLI (Loi de) : voir l'article « les Probabilités », page 437.

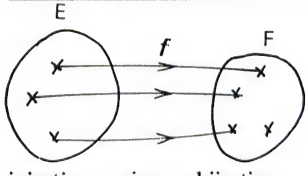
BIJECTION ◆ pages 20, 34 et 35, 80, 262, 277 et 278, 298, 375.

◆ Lire d'abord le mot Application.

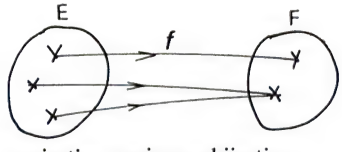
Soient deux ensembles E et F et une application f de l'ensemble E vers l'ensemble F. f est une bijection (ou est bijective) si :
— à tout élément de E correspond, par f, un élément de F (définition d'une application) ;
— à des éléments distincts de E correspondent, par f, des éléments distincts de F (c'est la définition de l'injection) ;
— tout élément de F provient, par f, d'au moins un élément de E (c'est la définition de la surjection).



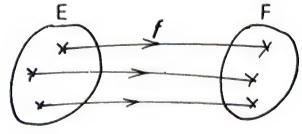
f: application



injection, mais pas bijection



surjection, mais pas bijection



bijection

L'application f de E vers F est bijective si elle est à la fois injective et surjective. Ces trois conditions réunies équivalent à la suivante :

Tout élément de F provient, par f, d'un élément de E et d'un seul.

Par exemple, l'application f de $\mathbb{R}$ (ensemble des nombres réels) vers $\mathbb{R}$, qui au nombre réel x fait correspondre son double $2x$, est une bijection de $\mathbb{R}$ sur $\mathbb{R}$, car tout nombre réel a un double, deux nombres réels distincts ont des doubles distincts (injection) et tout nombre réel peut être considéré comme le double d'un nombre réel, sa moitié (surjection).
$f : x \longmapsto 2x$ est une bijection de $\mathbb{R}$ vers $\mathbb{R}$.

BILINÉAIRE : voir Forme, Application et la page 315.

BINAIRE (numération, relation) : voir Numération ou Relation

BINÔME pages 408, 488.

Quand la somme de deux monômes n'est pas un monôme, c'est un binôme.
Exemples : si a, b, x, y sont des nombres réels :

$axb + 2\ axb$ est un monôme ($= 3\ axb$) ;
$ax + b$, $a^2x^3 + by$, $ax^2 - by$ sont des binômes.

Binôme de Newton

Les formules :

$$(a + b)^2 = a^2 + 2\ ab + b^2$$
$$(a - b)^2 = a^2 - 2\ ab + b^2$$
$$(a + b)^3 = a^3 + 3\ a^2b + 3\ ab^2 + b^3$$

que l'on apprend sous le nom d'identités remarquables, sont en fait des cas particuliers d'une formule plus générale qui donne le développement de $(a + b)^n$ pour n entier naturel quelconque. Cette formule est connue sous le nom de formule du binôme de Newton. Elle est valable pour n'importe quels éléments a et b d'un anneau A (pour les lois $+$ et $\times$), à condition que les deux éléments a et b soient tels que $a \times b = b \times a$:

$$(a + b)^n = a^n + C_n^1\, a^{n-1} \times b + C_n^2\, a^{n-2} \times b^2 + \ldots$$
$$+ C_n^p\, a^{n-p} \times b^p + \ldots + C_p^{n-1}\, a \times b^{n-1} + b^n.$$

Les C_n^p sont des nombres entiers naturels, appelés coefficients du binôme (ils expriment le nombre de combinaisons de n éléments pris p à p ; on peut les calculer facilement grâce au triangle de Pascal).

BINOMIALE (Loi) : voir l'article « les Statistiques », pages 486 à 490.

BIPOINT

Souvent employé comme synonyme de couple de points.

BIRAPPORT pages 318, 331.

Birapport de quatre nombres

Si a, b, c, d sont quatre nombres réels deux à deux distincts, on appelle birapport de ces quatre nombres, dans cet ordre, le nombre égal au quotient du rapport $\frac{a-c}{a-d}$ par le rapport $\frac{b-c}{b-d}$. On note ce nombre (a, b, c, d).

$$(a, b, c, d) = \frac{a-c}{a-d} : \frac{b-c}{b-d}.$$

Exemple

$$(1, 2, 3, 4) = \frac{1-3}{1-4} : \frac{2-3}{2-4} = \frac{-2}{-3} : \frac{-1}{-2} = \frac{2}{3} : \frac{1}{2} = \frac{4}{3}.$$

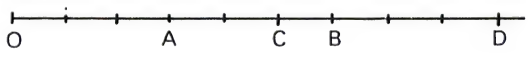
Birapport de quatre points situés sur un axe

Si A, B, C, D sont quatre points situés sur le même axe Ox, on appelle birapport de ces quatre points le nombre réel, noté (A, B, C, D), égal au birapport de leurs abscisses :

$$(A, B, C, D) = \frac{x_A - x_C}{x_A - x_D} : \frac{x_B - x_C}{x_B - x_D}$$

(ce birapport est indépendant de l'origine choisie sur l'axe, de l'orientation, de l'unité).
Si le birapport de quatre points A, B, C, D est égal à — 1, on dit que les quatre points forment une division harmonique.

Exemple

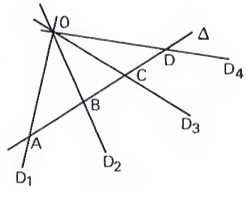


$$(A, B, C, D) = \frac{3-5}{3-9} : \frac{6-5}{6-9} = \frac{-2}{-6} : \frac{+1}{-3} = \frac{2}{6} : -\frac{1}{3} = -1.$$

Les quatre points A, B, C, D forment une division harmonique.

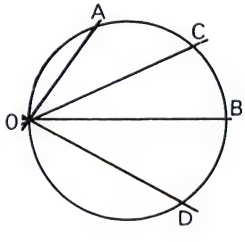
Birapport de quatre droites concourantes

Soient quatre droites concourantes D_1, D_2, D_3, D_4 et A, B, C, D leurs points d'intersection avec une droite quelconque Δ. Le birapport des quatre points A, B, C, D est le même, quelle que soit la droite Δ choisie. Ce birapport est appelé birapport des quatre droites D_1, D_2, D_3, D_4 et noté (D_1, D_2, D_3, D_4).
Si le birapport est égal à — 1, on dit que les quatre droites forment un faisceau harmonique.



Birapport de quatre points sur un cercle ou une conique

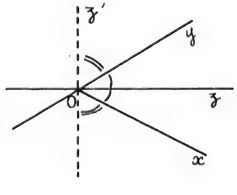
Si A, B, C, D sont quatre points situés sur le même cercle (ou la même conique) et si O est un point de cercle (ou de la conique), le birapport des quatre droites joignant O à A, B, C, D est indépendant du point O choisi ; c'est ce birapport qu'on appelle birapport des quatre points A, B, C, D.
Si le birapport des quatre points est égal à — 1, on dit qu'ils forment un quadrangle harmonique.



BISSECTRICE

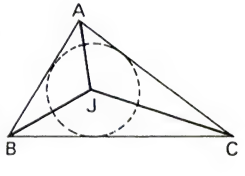
Bissectrice d'un angle géométrique

On appelle bissectrice de l'angle géométrique xOy toute droite Oz qui partage l'angle xOy en deux angles égaux. Il existe deux telles droites Oz et Oz′, qui sont perpendiculaires ; Oz est appelée bissectrice intérieure, Oz′ bissectrice extérieure.



Bissectrices d'un triangle

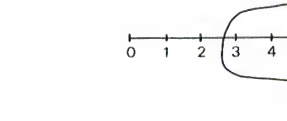
Les bissectrices des trois angles d'un triangle ABC sont appelées bissectrices du triangle. Les trois bissectrices intérieures sont concourantes en un point I qui est le centre du cercle inscrit dans le triangle.



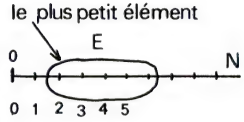
BON ORDRE pages 209, 282, 330, 377, 394.

L'ensemble des entiers naturels ℕ est muni d'une relation d'ordre, notée ⩽ (inférieur ou égal à) (l'entier p est inférieur ou égal à l'entier q, s'il existe un entier r tel que q soit égal à $p + r$).
N'importe quelle partie A de ℕ admet un élément qui est plus petit que tous les éléments de A :

Si $A = \{3, 4, 5, 6, \ldots, n\}$,
le plus petit élément de A est 3.



L'ensemble des nombres réels $\mathbb{R}$ est aussi muni d'une relation d'ordre, notée aussi $\leqslant$ (le réel x est inférieur ou égal au réel y si la différence $y - x$ est positive).
Une partie de $\mathbb{R}$ constituée par un intervalle fermé $[a, b]$ ou un intervalle semi-ouvert $[a, b[$ ou $[a, +\infty[$ admet bien un plus petit élément : a, mais pas une partie ouverte $]a, b[$ (car, si on considère un nombre x quelconque de $]a, b[$, aussi proche soit-il de a, il existe toujours un nombre réel y qui soit compris entre a et x, donc qui soit inférieur à x).



On dit que $\mathbb{N}$ est bien ordonné, ou encore que la relation d'ordre de $\mathbb{N}$ est une relation de bon ordre, tandis que la relation d'ordre de $\mathbb{R}$ n'est pas une relation de bon ordre.

De façon générale, une relation d'ordre dans un ensemble E est une relation de *bon ordre* si toute partie non vide de E admet un plus petit élément.

BOOLE (Algèbre de) : voir Algèbre de Boole

BORNÉ pages 105 à 107, 113, 117, 533, 534.

Ensemble borné

Dans l'ensemble $\mathbb{N}$ des entiers positifs, on a défini une relation « est inférieur ou égal à ». Si on considère le sous-ensemble de $\mathbb{N}$ constitué des trois nombres 3, 4, 5 : $A = \{ 3, 4, 5 \}$, tous les éléments de A sont inférieurs ou égaux à 5 et ils sont tous supérieurs ou égaux à 3 ; la partie A de $\mathbb{N}$ est bornée.
$\mathbb{N}$ n'est pas borné, car on ne peut trouver un entier p tel que tous les autres lui soient inférieurs.
Soit E un ensemble sur lequel est définie une relation d'ordre (notée $\leqslant$) ; une partie de A est dite bornée s'il existe deux éléments m et M de E tels que tout élément x de A soit à la fois inférieur ou égal à M et supérieur ou égal à m :

$$\forall\, x \in A, \qquad m \leqslant x \leqslant M.$$

Tout intervalle fini de $\mathbb{R}$ (ensemble des nombres réels) est borné.

Application bornée ◆

◆ Lire d'abord Application.

Une application f d'un ensemble E vers un ensemble F muni d'une relation d'ordre est dite bornée si l'image par f de E est une partie bornée de F.
Par exemple, l'application f de l'intervalle $[0, 1]$ dans $\mathbb{R}$ (ensemble des nombres réels) qui à tout nombre réel x de $[0, 1]$ fait correspondre le nombre réel $f(x) = 2x + 1$ est bornée, car si x est compris entre 0 et 1, alors $f(x)$, c'est-à-dire $2x + 1$, est compris entre 1 et 3.

BOULE ◆ page 248.

◆ Lire d'abord Distance.

Soient E un ensemble muni d'une distance d, x un point de E et r un nombre réel positif.
On appelle boule ouverte de centre x et de rayon r et on note $B(x, r)$ l'ensemble des éléments y de E dont la distance à x, $d(x, y)$, est strictement inférieure à r :

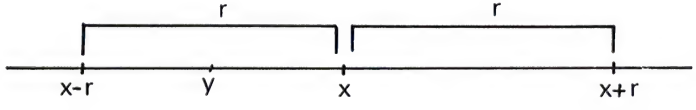
$$B(x, r) = \{ y \mid y \in E, d(x, y) < r \}.$$

On appelle boule fermée de centre x et de rayon r l'ensemble des éléments y de E dont la distance $d(x, y)$ à x est inférieure ou égale à r.

On la note quelquefois $\overline{B}(x, r)$:

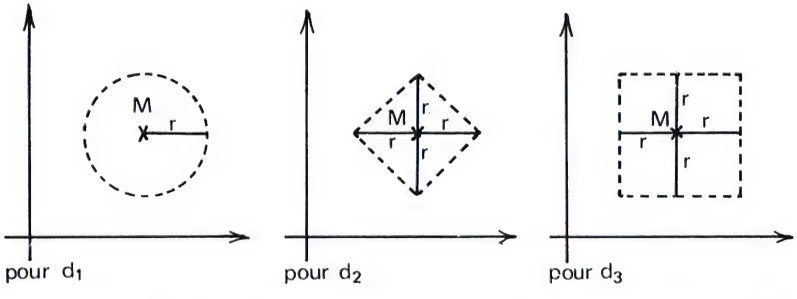
$$\overline{B}(x, r) = \{ y \mid y \in E, d(x, y) \leqslant r \}.$$

Par exemple, dans l'ensemble $\mathbb{R}$ des nombres réels sur lequel on a défini une distance d (la distance de deux nombres réels étant la différence entre le plus grand et le plus petit), la boule ouverte de centre x et de rayon r est l'ensemble des nombres réels y dont la distance à x est inférieure à r. C'est donc l'ensemble des nombres réels, y compris strictement entre $x - r$ et $x + r$; c'est l'intervalle ouvert $]x - r, x + r[$.

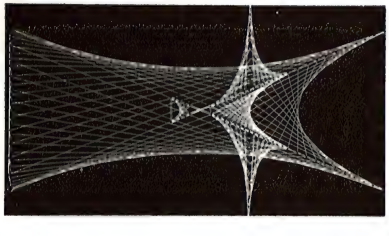


La boule fermée de centre x et de rayon r est l'ensemble des nombres réels y dont la distance à x est inférieure ou égale à r ; c'est donc l'intervalle fermé $[x - r, x + r]$.

L'ensemble $\mathbb{R}^2$ des couples (x, y), où x et y sont réels, peut être muni d'une distance de trois façons différentes ; pour chacune des distances possibles, les boules ouvertes ont la forme suivante : $(M = (x, y))$.



Les contours sont en pointillés, car ils ne font pas partie de la boule ouverte (mais ils font partie de la boule fermée).



C

CALCUL VECTORIEL : voir Algèbre linéaire

CALCUL MATRICIEL : voir l'article « l'Algèbre linéaire », pages 38 à 43.

CANONIQUE◆

◆ Lire d'abord Application.

Se dit d'un objet mathématique attaché à certaines structures de façon privilégiée.

Injection canonique

Si A est une partie d'un ensemble E, l'application qui à chaque élément x de A associe le même élément x considéré comme élément de E (c'est donc une application de A dans E) est appelée injection canonique de A dans E.

Base canonique ◆

◆ Lire d'abord l'article « l'Algèbre linéaire » : base d'un espace vectoriel, page 65.

$\mathbb{R}^2$, ensemble des couples (x, y) d'éléments de $\mathbb{R}$, est un espace vectoriel de dimension 2, dont une base simple est constituée par l'ensemble des deux couples (0, 1) et (1, 0). Cette base est appelée base canonique de $\mathbb{R}^2$. De même, la base constituée par les trois triplets (1, 0, 0), (0, 1, 0) et (0, 0, 1) est appelée base canonique de l'espace vectoriel $\mathbb{R}^3$ (ensemble des triplets (x, y, z) d'éléments de $\mathbb{R}$).

Application canonique ◆

◆ Lire d'abord Équivalence (relation d').

Si E est un ensemble sur lequel est définie une relation d'équivalence $\mathcal{R}$, l'application qui à chaque élément x de E associe sa classe d'équivalence $\mathrm{Cl}(x)$ est appelée application canonique de E. C'est une application de E dans l'ensemble quotient $E/\mathcal{R}$.

CARACTÉRISTIQUE pages 52 et 53, 348.

Fonction caractéristique ◆

◆ Lire d'abord Application et Complémentaire.

Si E est un ensemble et A un sous-ensemble de E, la fonction caractéristique de A est l'application de A dans l'ensemble $\{0, 1\}$ qui prend :

— la valeur 1 pour tout x de A ;

— la valeur 0 pour tout x appartenant au complémentaire de A.

On la note en général χ_A

$$\forall x \in A, \quad \chi_A(x) = 1$$

(lire : pour tout x appartenant à A, la valeur khi$_A$ de x est égale à 1).

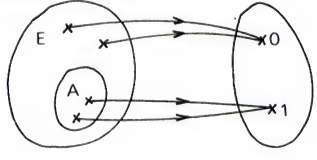
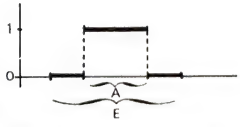


Diagramme sagittal de χ_A

$$\forall x \in \complement_E A, \quad \chi_A(x) = 0$$

(lire : pour tout x appartenant au complémentaire de A dans E, la valeur khi$_A$ est 0).



Représentation cartésienne de χ_A dans le cas où A et E sont des intervalles.

Si A et B sont deux parties du même ensemble E, si χ_A et χ_B sont leurs fonctions caractéristiques respectives, la fonction caractéristique de l'intersection $A \cap B$ est égale au produit des fonctions χ_A et χ_B :

$$\chi_{A \cap B} = \chi_A \times \chi_B$$

La fonction caractéristique de la réunion est égale à la somme de χ_A et χ_B moins le produit $\chi_A \times \chi_B$:

$$\chi_{A \cup B} = \chi_A + \chi_B - \chi_A \cdot \chi_B.$$

Caractéristique d'un anneau ◆

◆ Il est conseillé de lire l'article « l'Algèbre », page 46.

Soit E un ensemble qui, muni des deux lois de composition interne notées + et ×, a structure d'anneau. Quel que soit l'élément x de E et quel que soit le nombre entier relatif n, on définit nx comme étant :
— 0 si $n = 0$;
— la somme de n termes égaux à x si n est un entier positif ;
— la somme de $-n$ termes égaux à $-x$ (symétrique de x pour l'addition) si n est négatif.
Si 0 est l'élément neutre de la loi +, ou bien il n'existe aucun entier naturel $m \neq 0$ tel que $m \,.\, x = 0$ pour tout x de E (0 est donc le seul entier naturel pour lequel $m \,.\, x = 0$), on dit que E est de caractéristique nulle, ou bien il existe un entier naturel m non nul, tel que $m \,.\, x = 0$ pour tout x de E, et alors il en existe une infinité (car, pour tout entier p, $(pm)x = p \,.\, (m \,.\, x) = p \,.\, 0 = 0$) ; donc, pour tous les multiples de m, $(pm)\, x = 0$. Dans ce cas, il existe un entier naturel minimal n tel que, pour tout x de E, $n \,.\, x = 0$. On dit que E est de caractéristique n.
$\mathbb{R}$ (l'ensemble des nombres réels), $\mathbb{C}$ (l'ensemble des nombres complexes) sont de caractéristique nulle.

$\mathbb{Z}/2\,◆ = \{ 0, 1 \}$ est de caractéristique deux car : $\begin{cases} 2 \,.\, 0 = 0 + 0 = 0 \\ 2 \,.\, 1 = 1 + 1 = 0. \end{cases}$

Table d'addition de $\mathbb{Z}/2$.

+	0	1
0	0	1
1	1	0

Si l'anneau E *est unitaire*, soit 1 l'élément neutre de la loi × et s'il existe des nombres entiers naturels tels que $m \,.\, 1 = 0$. Si n est le plus petit d'entre eux, alors la caractéristique de A est n.

Si l'anneau E *est un anneau intègre*, non réduit à $\{ 0 \}$, et si la caractéristique de E n'est pas nulle, cette caractéristique est un nombre premier (cela est vrai en particulier pour un corps).

Dans un anneau commutatif de caractéristique p ($p \neq 0$), où p est un nombre premier, la formule donnant l'expression développée de $(a + b)^p$ est la suivante :

$$(a + b)^p = a^p + b^p$$

car tous les coefficients apparaissant dans le développement de $(a + b)^p$ sont tous multiples de p ; les termes correspondants sont donc nuls.

Caractéristique d'un logarithme : voir le mot Logarithme.

Caractéristique (équation) d'une matrice : voir l'article « l'Algèbre linéaire ».

CARDINAL : voir Equipotent, l'article « les Ensembles », pages 264 à 270 et les pages 283 et 284, 298, 377, 378.

CARDIOÏDE : voir Cycloïde

CAROLL (Diagramme de) : voir Diagramme

CARRÉ (Square) : voir Polygones

CARRÉ LATIN : voir l'article « l'Algèbre », page 31.

CARTÉSIEN (Diagramme) : voir Diagramme et les pages 269, 274, 279 et 280.

CARTÉSIENNE (Coordonnées) : voir Coordonnées

CATÉGORIE

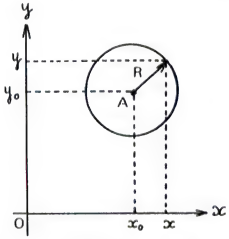
Notion nouvelle généralisant la notion d'ensemble.

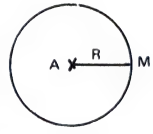
CENTRE : voir Boule, Cercle, Disque, Sphère

CERCLE

Dans le plan euclidien (c'est-à-dire plan muni de la distance euclidienne ◆), on appelle cercle de centre A et de rayon R l'ensemble des points M du plan situés à une même distance R (R étant un nombre réel positif) d'un point A.

◆ Voir Distance.





Si le plan est muni d'un repère orthonormé Ox, Oy, *l'équation du cercle*, de centre A, de coordonnées (x_0, y_0) et de rayon R est :

$$(x - x_0)^2 + (y - y_0)^2 = R^2.$$

En particulier, si le centre du cercle est l'origine O du repère, l'équation s'écrit :

$$x^2 + y^2 = R^2$$

Longueur du cercle : $2\pi R$.

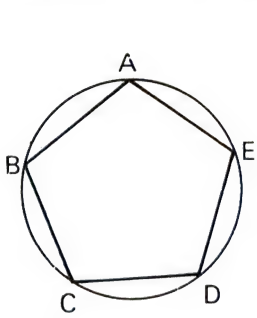
Surface du cercle (que l'on emploie par abus de langage, car on devrait parler de surface du disque, c'est-à-dire de la surface intérieure au cercle) : πR^2.

CHASLES (Relation de) : voir Algébrique (mesure) et la page 98.

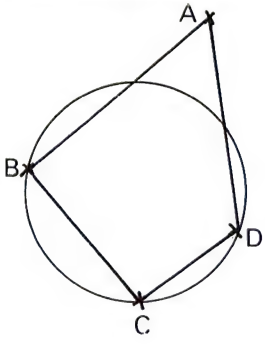
CIRCONSCRIT (Cercle)

— *Cercle circonscrit à un triangle :* cercle passant par les trois sommets du triangle ; le centre de ce cercle est le point d'intersection des trois médiatrices du triangle.

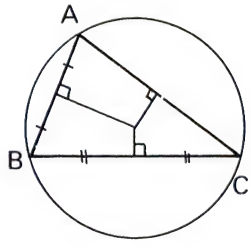
— *Cercle circonscrit à un polygone :* cercle passant par tous les sommets du polygone. Certains polygones admettent un cercle circonscrit, d'autres non. Ceux qui en admettent un sont dits inscriptibles (les triangles, les polygones réguliers sont des polygones inscriptibles).



polygone inscriptible



polygone non inscriptible



cercle circonscrit à un triangle

CLASSE (d'équivalence) : voir Equivalence et les pages 262, 298, 377, 458.

CLASSE STATISTIQUE : voir l'article « les Statistiques », page 460.

CLOS page 375.

Le polynôme $x^2 + 1 = 0$ n'a pas de racines dans l'ensemble $\mathbb{R}$ des nombres réels ; par contre, il a deux racines complexes : $+ i$ et $- i$. S'il existe des polynômes à coefficients réels n'ayant pas de racines réelles, par contre, tout polynôme à coefficients réels ou complexes a des racines complexes. On dit que le corps $\mathbb{C}$ (ensemble des nombres complexes) est algébriquement clos.

COCYCLIQUE

Des points sont dits cocycliques s'ils appartiennent à un même cercle. Deux points, trois points non alignés sont toujours cocycliques.

COEFFICIENT DIRECTEUR D'UNE DROITE pages 239 et 240.

Dans un repère Ox, Oy, l'équation d'une droite non parallèle à Oy est de la forme :

$y = ax + b$, où a et b sont deux nombres réels.

Le nombre a est appelé coefficient directeur de la droite.

Si le repère Ox, Oy est orthonormé, ce coefficient directeur est encore appelé pente de la droite.

COLINÉAIRE : voir l'article « l'Algèbre linéaire », pages 58 et 63.

COLLECTION : voir l'article « les Ensembles », pages 261, 272.

COMBINAISON page 66.

Soit un ensemble E de quatre éléments $\{a, b, c, d\}$; le nombre de sous-ensembles possibles est de $2^4 = 16$.

Sous-ensembles à 0 élément : $\varnothing$ (ensemble vide).
Sous-ensembles à 1 élément : $\{a\}, \{b\}, \{c\}, \{d\}$.
Sous-ensembles à 2 éléments : $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$.
Sous-ensembles à 3 éléments : $\{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}$.
Sous-ensemble à 4 éléments : $\{a, b, c, d\}$.

Parmi toutes ces parties, celles qui ont deux éléments ◆ sont $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$; il y en a six. On peut remarquer que le nombre des parties à deux éléments est égal à la moitié du nombre des arrangements sans répétition des quatre éléments pris deux par deux, car, à partir de deux éléments, a et b par exemple, on peut former deux arrangements ab et ba, mais une seule partie $\{a, b\}$.

◆ Lire d'abord Arrangement.

Il y a quatre parties de E qui ont trois éléments : ce nombre est égal au nombre, divisé par 6, des arrangements sans répétition (24) des quatre éléments de E pris 3 par 3, car, à partir de trois éléments a, b, c, on peut former une seule partie: $\{a, b, c\}$, mais six arrangements : abc, acb, bca, bac, cab, cba ; puis six arrangements pour $\{a, b, d\}$, etc.

De façon générale, si E est un ensemble à n éléments et si p est un entier naturel compris entre 0 et n, on appelle combinaison des n éléments de E pris p à p toute partie de E qui contient p éléments. Le nombre de ces combinaisons est noté C_n^p◆ ; il est égal au nombre d'arrangements A_n^p divisé par $p!$◆ (car le même groupe de p éléments permet de former $p!$ arrangements distincts, mais une seule partie de E)

◆ On le note aussi quelquefois $\binom{n}{p}$.

◆ Voir le mot Factorielle.

$$C_n^p = \frac{n(n-1)(n-2)\ldots(n-p+1)}{p!} = \frac{A_n^p}{p!}.$$

Exercice

Combien peut-on faire de combinaisons de 8 cartes prises parmi un jeu de 32 cartes ?

$$C_n^p = C_{32}^8 = \frac{32 \times 31 \times 30 \times \ldots \times 25}{1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8}.$$

On écrit aussi C_n^p sous la forme, plus symétrique, suivante :

$$C_n^p = \frac{n!}{p!\,(n-p)!}$$

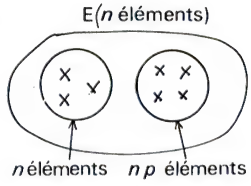
(obtenue en multipliant la première expression au numérateur et au dénominateur par $(n-p)!$).

● $C_n^0 = C_n^n = 1$ (car il y a une seule partie à 0 élément : la partie $\varnothing$ (vide), et une seule à n éléments : E lui-même).

● $C_n^p = C_n^{n-p}$ (car il y a autant de parties à p éléments que de parties à $n - p$ éléments).

● $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$.

Cette dernière formule permet d'établir le triangle de Pascal.



COMMUTATIF, COMMUTATIVITÉ◆ pages 21 et 22, 56, 73 à 75, 267, 270.

◆ Lire d'abord Composition (loi de) et l'article « l'Algèbre », page 21.

La loi de composition interne, notée $\divideontimes$, définie sur l'ensemble E, est commutative si, quel que soit le couple (a, b) d'éléments de E, le composé de a avec b : $a \divideontimes b$ et le composé de b avec a : $b \divideontimes a$ sont deux éléments égaux :

$$\forall\, a \in E,\ \forall\, b \in E \quad a \divideontimes b = b \divideontimes a$$

(quels que soient les éléments a et b de E, a étoile b est égal à b étoile a).

Par exemple, addition, multiplication sont deux lois de composition interne sur $\mathbb{N}$ (entiers naturels positifs) commutatives ; la soustraction, loi de composition interne sur $\mathbb{Z}$, n'est pas commutative, car il existe des couples pour lesquels $a - b$ est distinct de $b - a$ (ainsi $3 - 2 = +1$; $2 - 3 = -1$).

Groupe commutatif ou abélien ◆

◆ Il est conseillé de lire l'article « l'Algèbre », page 28.

Si un ensemble E a structure de groupe pour la loi de composition interne $\divideontimes$, on dit que ce groupe est commutatif si la loi $\divideontimes$ est commutative (on dit aussi qu'il est abélien).

Anneau commutatif ou corps ◆

◆ Il est conseillé de lire l'article « l'Algèbre », page 46 et page 53.

Si un ensemble E muni de deux lois de composition interne (notées $+$ et $\times$) a une structure d'anneau (ou de corps), on dit que c'est un anneau (ou un corps) commutatif si la loi notée $\times$ est commutative (la loi notée $+$ est commutative ◆)

◆ Voir Anneau.

COMPACT◆

◆ Lire d'abord Topologie, Séparé, Recouvrement.

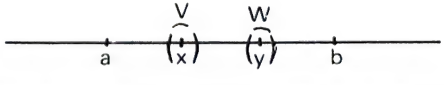
Un espace topologique E est compact s'il satisfait aux deux propriétés suivantes :

1) il est séparé ;

2) de tout recouvrement ouvert de E, on peut extraire un recouvrement fini.

Les intervalles fermés bornés $[a, b]$ de $\mathbb{R}$, ensemble des nombres réels, sont des ensembles compacts, car :

1) $[a, b]$ est séparé ; quels que soient les points distincts x et y de cet intervalle, on peut toujours trouver un voisinage de x et un voisinage de y qui ne se rencontrent pas ;



2) De tout recouvrement ouvert de $[a, b]$, on peut extraire un recouvrement fini (cela se démontre).

$\mathbb{R}$ n'est pas compact. La droite numérique achevée $\overline{\mathbb{R}}$ est compacte.

COMPARABLE

Si E est un ensemble muni d'une relation d'ordre notée $\leqslant$, deux éléments x et y de E sont comparables si l'une au moins des deux propositions « x est inférieur à y » ou « y est inférieur à x » est vraie.
Dans le cas d'une relation d'ordre total, deux éléments quelconques de E sont toujours comparables ; dans le cas d'une relation d'ordre partiel, il existe des couples d'éléments de E non comparables.

COMPATIBLE

— Si E est un ensemble sur lequel sont définies une relation d'équivalence $\mathcal{R}$ et une loi de composition interne notée $\divideontimes$, on dit que $\mathcal{R}$ est compatible avec la loi $\divideontimes$ si, lorsque x est équivalent à x' et y à y', alors le composé $x \divideontimes y$ est équivalent au composé $x' \divideontimes y'$.

$$\left.\begin{array}{l} x \mathcal{R} x' \\ y \mathcal{R} y' \end{array}\right| \Longrightarrow (x \divideontimes y) \mathcal{R} (x' \divideontimes y')$$

Par exemple, soit dans $\mathbb{Z}$ (ensemble de nombres entiers relatifs, négatifs et positifs) la relation $\mathcal{R}$ définie par « x a même reste dans la division par 2 que y ». Cette relation d'équivalence est compatible avec l'addition, car, si x a même reste que x' et y même reste que y' dans la division par 2, $x + y$ a même reste que $x' + y'$ (dans la division par 2). Ainsi 3 et 5 ont même reste dans la division par 2 (ce reste est 1). 2 et 6 ont même reste dans la division par 2 (ce reste est 0).
Alors : 3 + 2 et 5 + 6 ont même reste dans la division par 2 (c'est-à-dire 1).

— Si E est un ensemble muni d'une loi de composition externe $x \longmapsto \lambda \cdot x$ (dont les opérateurs appartiennent à un ensemble Ω) et d'une relation d'équivalence $\mathcal{R}$, $\mathcal{R}$ est dite compatible avec la loi externe si, lorsque x est équivalent à x', alors, quel que soit l'élément λ de Ω, $\lambda \cdot x$ est équivalent à $\lambda \cdot x'$:

$$x \mathcal{R} x' \Longrightarrow \forall \lambda \, (\lambda \cdot x) \mathcal{R} (\lambda \cdot x')$$

COMPLÉMENTAIRE pages 363, 364.

A étant une partie de l'ensemble E, on appelle complémentaire de A dans E et on note $\complement_E A$ ◆ l'ensemble des éléments de E qui n'appartiennent pas à A :

$$\complement_E A = \{ x \mid x \in E \text{ et } x \notin A \}$$

(on lit : le complémentaire du sous-ensemble A dans l'ensemble E est l'ensemble des x qui appartiennent à E et qui n'appartiennent pas à A [voir figure ci-contre]).

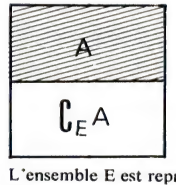
● L'intersection de A avec son complémentaire est vide (puisqu'il n'y a aucun élément qui puisse à la fois appartenir à A et ne pas appartenir à A) :

$$A \cap \complement_E A = \varnothing.$$

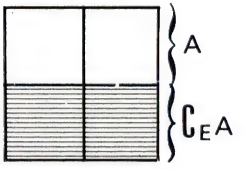
● La réunion de A et de son complémentaire est égale à E :

$$A \cup \complement_E A = E.$$

◆ On lit « complémentaire de A dans E ».
Si $E = \{ 1, 2, 3, 4 \}$,
$A = \{ 1, 2 \}$,
$\complement_E A = \{ 3, 4 \}$.



L'ensemble E est représenté par la surface intérieure au carré.



● Si A est vide, son complémentaire dans E est E :

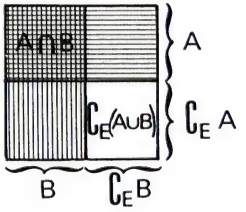
$$\complement_E \varnothing = E.$$

● Si A est E lui-même, son complémentaire (dans E) est vide :

$$\complement_E E = \varnothing.$$

● Le complémentaire (dans E) du complémentaire (dans E) de A est égal à A :

$$\complement_E(\complement_E A) = A.$$



● Si A et B sont deux parties du même ensemble E, leur intersection est une partie de E dont le complémentaire (dans E) est la réunion du complémentaire (dans E) de A et du complémentaire (dans E) de B (voir figure ci-contre) :

$$\complement_E(A \cap B) = (\complement_E A) \cup (\complement_E B).$$

● Si A et B sont deux parties de E, leur réunion A ∪ B est une partie de E dont le complémentaire (dans E) est l'intersection du complémentaire de A avec le complémentaire de B :

$$\complement_E(A \cup B) = (\complement_E A) \cap (\complement_E B).$$

COMPLET (Espace) : voir Suite de Cauchy

COMPLEXE (Nombre) pages 162, 229, 286, 336, 374 et 375.

Le carré d'un nombre réel est toujours positif ◆. Pour résoudre certaines équations, les mathématiciens ont été amenés ◆ à introduire un nombre « imaginaire » noté i dont le carré est égal à -1 (tout nombre réel négatif peut alors être considéré comme le carré d'un nombre imaginaire : par exemple, -4 est le carré de $2i$, $(2i)^2$). Les nombres complexes sont de la forme $a + bi$ où a et b sont des nombres réels et i ce nombre imaginaire.

◆ Exemples : $(-2)^2 = +4$; $2^2 = +4$.

◆ Voir l'article « les Applications des mathématiques », page 186.

Pour additionner, soustraire ou multiplier les nombres complexes, on procède exactement comme pour additionner, soustraire ou multiplier des nombres réels en tenant compte (chaque fois que cela se présente) du fait que $i^2 = i \times i = -1$.

Exemple :

$$\begin{aligned}(2 + 3i) \times (-1 + 4i) &= 2 \times (-1) + 2 \times (4i) + 3i \times (-1) + (3i) \times (4i)\\ &= -2 + 8i - 3i + 12i^2\\ &= -2 + 8i - 3i - 12\\ &= -14 + 5i\end{aligned}$$

L'ensemble des nombres complexes est noté $\mathbb{C}$. Les nombres réels appartiennent à $\mathbb{C}$ ◆.

◆ Voir l'article « les Nombres ».

COMPOSANTE page 72.

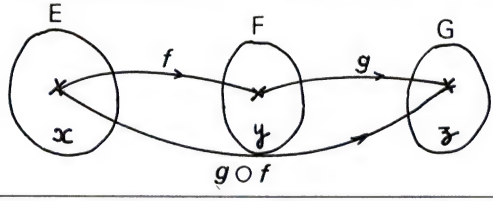
Dans un couple (x, y), l'élément x est appelé première composante du couple et l'élément y deuxième composante.

COMPOSÉ page 276.

Le résultat de l'addition de deux nombres entiers x et y est appelé somme de ces deux nombres et noté $x + y$; le résultat de la multiplication de l'entier x par l'entier y est appelé produit de ces deux nombres et noté $x \times y$ ou $x \cdot y$ ou xy. De façon générale, si un ensemble E est muni d'une loi de composition interne, dont le symbole est noté $\divideontimes$ (par exemple), le résultat de la composition de x et de y est appelé le composé des deux éléments x et y et noté $x \divideontimes y$.

COMPOSITION (des applications) COMPOSÉE (Application) pages 248, 97, 216, 224.

Soient E, F, G trois ensembles, f une application de E vers F et g une application de F vers G.

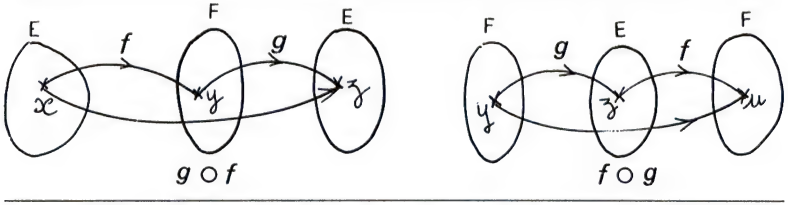


A chaque élément x de E correspond par f un élément y de F : $y = f(x)$; à cet élément y de F, on peut faire correspondre par g un élément z de G : $z = g(y)$; ainsi, par ce procédé, à chaque élément x de E on a pu faire correspondre un élément z de G ; on a donc défini une application de E dans G, que l'on appelle application composée de l'application f par l'application g et que l'on note $g \circ f$ (on lit : g rond f) :

$$z = g(y) = g\,[f(x)] = g \circ f(x)$$

Pour que l'on puisse composer l'application f et l'application g : $g \circ f$, il faut que l'ensemble d'arrivée F de f, f : E ⟶ F, soit l'ensemble de départ de g, g : F ⟶ G (pour qu'on puisse parler des images par g des éléments y) ; il ne sera donc possible de composer g et f ($f \circ g$) que si l'ensemble d'arrivée de g, G, est égal à l'ensemble de départ de f, E.

Dans le cas où E = G, on pourra donc parler de $g \circ f$ ainsi que de $f \circ g$; $g \circ f$ est une application de E dans E, tandis que $f \circ g$ est une application de F dans F. Les deux applications $f \circ g$ et $g \circ f$ sont donc deux applications distinctes. Même dans le cas où E = F, les deux applications $f \circ g$ et $g \circ f$ ne sont pas égales en général.



Exemple

Soient f, g deux applications de $\mathbb{R}$ (ensemble des nombres réels) dans $\mathbb{R}$.

$f : x \longmapsto 2x + 3$ (qui à x fait correspondre $2x + 3$)

$g : x \longmapsto x^2$ (qui à x fait correspondre x^2)

$$g \circ f : \left.\begin{array}{l} x \overset{f}{\longmapsto} 2x + 3 = y \\ y \overset{g}{\longmapsto} y^2 = z \\ g \circ f(x) = (2x + 3)^2 \end{array}\right\} z = (2x + 3)^2$$

$$f \circ g : \left.\begin{array}{l} x \overset{g}{\longmapsto} x^2 = y \\ y \overset{f}{\longmapsto} 2y + 3 = z \\ f \circ g\ (x) = 2x^2 + 3 \end{array}\right\} z = 2x^2 + 3$$

$f \circ g\ (x)$ et $g \circ f\ (x)$ sont distincts (sauf pour deux valeurs de x) donc $f \circ g \neq g \circ f$.

Par contre, si l'application f est l'application de $\mathbb{R}$ dans $\mathbb{R}$ qui à x fait correspondre son carré x^2 et si l'application g est l'application de $\mathbb{R}$ dans $\mathbb{R}$ qui à x fait correspondre son cube x^3 :

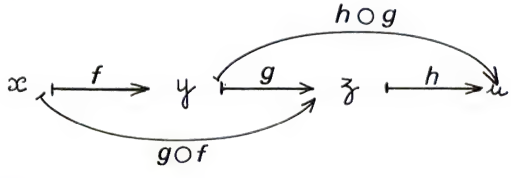
$$f \circ g : x \overset{g}{\longmapsto} x^3 \overset{f}{\longmapsto} (x^3)^2 = x^6$$

$$g \circ f : x \overset{f}{\longmapsto} x^2 \overset{g}{\longmapsto} (x^2)^3 = x^6$$

Quel que soit l'élément x de $\mathbb{R}$, on a $f \circ g\ (x) = g \circ f\ (x)$; les deux applications $f \circ g$ et $g \circ f$ sont égales.

Si f, g, h sont trois applications, f de E dans F, g de F dans G et h de G dans H, on peut composer f avec g : $g \circ f$, puis cette nouvelle application avec h ; on obtient donc l'application $h \circ (g \circ f)$; ou bien on peut composer f avec la composée $h \circ g$, obtenant ainsi $(h \circ g) \circ f$. Ces deux applications $h \circ (g \circ f)$ et $(h \circ g) \circ f$ sont égales (la loi de composition des applications est associative) ; on écrit donc tout simplement $h \circ g \circ f$.

$$[(h \circ g) \circ f]\ (x) = u = [h \circ (g \circ f)]\ (x)$$



COMPOSITION (Loi de) ◆ pages 17 à 27, 201, 408, 455.

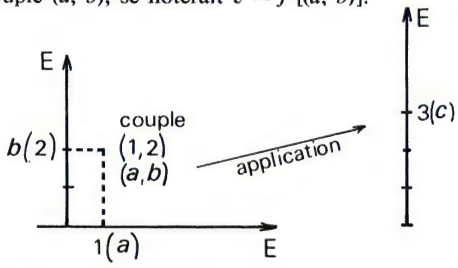
◆ Lire aussi l'article « l'Algèbre », page 17.

Loi de composition interne

Quel que soit le couple (a, b) de nombres entiers positifs (naturels), nous savons lui associer un troisième entier naturel que l'on appelle la somme de ces deux entiers et que l'on note $a + b$; le procédé nous permettant de calculer cette somme s'appelle l'addition.

De façon générale, une loi de composition interne sur un ensemble E est une application qui, à tout couple (a, b) d'éléments de E, fait correspondre un troisième élément de E : c (c'est donc une application du produit cartésien E $\times$ E dans E)

Si on notait f cette application (comme on le fait dans le cas général), c, étant l'image du couple (a, b), se noterait $c = f[(a, b)]$.



La loi de composition interne sur l'ensemble E est l'application du produit cartésien E × E (E croix E) vers E. Ici, la loi de composition est l'addition.

On préfère toutefois symboliser la loi de composition par un signe : ✲ (T, +, ×, ○, ...) et noter $c = a \text{ ✲ } b$ (lire : a étoile b), comme, pour l'addition, on note $c = a + b$; l'élément c est appelé le composé de a et de b par la loi ✲. Dans le cas de la loi addition (+), $a + b$ est appelé somme de a et de b ; dans le cas de la multiplication (×), $a \times b$ est appelé produit de a par b.

Exemples de lois de composition interne

— Addition, multiplication dans les ensembles $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ (entiers naturels $\mathbb{N}$, entiers relatifs $\mathbb{Z}$, nombres rationnels $\mathbb{Q}$, nombres réels $\mathbb{R}$, nombres complexes $\mathbb{C}$).
— Réunion, intersection dans l'ensemble des parties d'un ensemble.
— Loi ○ de composition des applications dans l'ensemble des applications d'un ensemble dans lui-même.

Remarque

La soustraction n'est pas une loi de composition interne dans l'ensemble des entiers positifs (naturels) $\mathbb{N}$, car il existe des couples (a, b) pour lesquels la différence $a - b$ n'existe pas dans $\mathbb{N}$ (quand b est supérieur à a, il n'y a pas de résultat positif). Par contre, c'est bien une loi de composition interne dans $\mathbb{Z}$ (ensemble des entiers relatifs positifs et négatifs), car, pour tout couple (a, b) d'éléments de $\mathbb{Z}$, la différence $a - b$ existe dans $\mathbb{Z}$. Si b est supérieur à a, le résultat est négatif ($5 - 7 = -2$).

Loi de composition externe

Quel que soit le polynôme ◆ P à coefficients dans $\mathbb{Z}$ (par exemple) et quel que soit l'entier naturel n, on sait associer à ce couple (n, P) un nouveau polynôme, que l'on note $n \,.\, P$ (dont tous les coefficients sont multipliés par n) et qui est encore un polynôme à coefficients dans $\mathbb{Z}$. On vient de définir ici une loi de composition qui met en jeu un nombre et un polynôme (donc deux éléments appartenant à deux ensembles distincts ; c'est pourquoi on la qualifie d'externe) et dont le résultat est un polynôme.
Le nombre n est appelé opérateur, car on peut considérer qu'il agit sur le polynôme P pour le transformer en un autre polynôme $n.P$.
Une loi de composition externe définie sur un ensemble E et dont les opérateurs sont dans un ensemble Ω est une application qui à tout couple (α, x) (α étant élément de Ω et x élément de E) fait correspondre un élément de E (que l'on note souvent $\alpha.x$).
Une loi de composition externe est donc une application du produit cartésien $\Omega \times E$ dans E.

◆ Un polynôme (à une seule variable x) est une expression algébrique de la forme
$a_0 + a_1x + a_2x^2 + \ldots a_n x^n$,
où les coefficients $a_0, a_1, a_2, \ldots, a_n$ sont des nombres réels, ou complexes, ou entiers et où n est un nombre entier positif.

COMPRÉHENSION page 260.

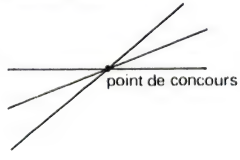
Un ensemble est défini en compréhension s'il est défini à l'aide d'une propriété que possèdent tous ses éléments et qu'ils sont les seuls à posséder.
Par exemple, la phrase « les nombres entiers positifs ou nuls multiples de 3 et plus petits que 20 » définit l'ensemble $\{ 0, 3, 6, 9, 12, 15, 18 \}$ en compréhension ; on note aussi :

$$\{ x \mid x \in \mathbb{N},\ x \text{ multiple de 3 et inférieur à } 20 \} \blacklozenge$$

◆ Lire :
l'ensemble des x dont la propriété est d'appartenir à l'ensemble des entiers naturels, etc.

CONCOURANT

Plusieurs droites sont dites concourantes si elles passent toutes par le même point (qui est alors appelé point de concours de ces droites).



CONDITION

Condition nécessaire

Pour qu'un nombre entier positif soit premier, il faut qu'il soit impair (car un nombre pair différent de 2 étant divisible par 2, il n'est pas un nombre premier).
Cette phrase peut s'écrire:

$$x \text{ premier} \Longrightarrow x \text{ impair.}$$

On dit que la condition « x est impair » est une condition nécessaire pour que x soit premier.

Condition suffisante

La condition « x est impair » ne suffit pas pour que x soit premier (par exemple, 15 est impair mais non premier, puisque divisible à la fois par 3 et par 5).
Cette condition n'est pas suffisante, car on n'a pas :

$$x \text{ impair} \Longrightarrow x \text{ premier.}$$

Par contre, x étant un nombre entier positif, la condition : « la somme des chiffres de x est un multiple de 3 » est une condition suffisante pour que x soit divisible par 3, car on a l'implication :
Somme des chiffres de x multiple de 3 $\Longrightarrow$ x divisible par 3.

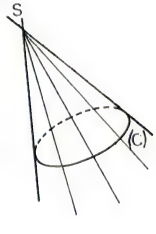
Condition nécessaire et suffisante

— Si x est divisible par 3, alors la somme de ses chiffres est multiple de 3.
— Si la somme des chiffres de x est multiple de 3, alors x est divisible par 3.

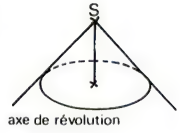
On dit que la condition : « la somme des chiffres de x est multiple de 3 » est une condition nécessaire et suffisante pour que x soit divisible par 3.

CÔNE

Ensemble de droites passant par un même point, appelé sommet du cône, et rencontrant une courbe (C) appelée directrice du cône.
Les droites sont appelées *génératrices* du cône.



Si la courbe (C) est un cercle et si le point S se trouve sur l'axe de ce cercle, le cône est dit *cône de révolution*. L'axe du cercle est appelé axe du cône.



CONGRUENCE (Modulo) page 50.

Si on fait la différence $x - y$ de deux entiers relatifs (positifs ou négatifs) x et y, ou bien cette différence est un *multiple* de 3, ou non. Si la différence est un multiple de 3, on dit que les deux entiers x et y sont congrus modulo 3. Par exemple, 7 et 4 sont congrus modulo 3, puisque $7 - 4 = 3$; 18 et 3 sont congrus modulo 3, puisque la différence est $18 - 3 = 15 = 3 \times 5$; 7 et 5 ne sont pas congrus modulo 3, car la différence $7 - 5 = 2$ n'est pas un multiple de 3.

De façon générale, si n est un entier strictement positif, on dit que deux entiers relatifs x et y sont congrus modulo n si leur différence est un multiple de n. On note :

$$x \equiv y \pmod{n}$$

et on lit : « x est congru à y modulo n ».

Cette relation, définie entre éléments de Z, ensemble des entiers relatifs, est appelée relation de congruence modulo n. C'est une relation d'équivalence, car elle est :

— réflexive : x est congru à lui-même modulo n (puisque la différence $x - x$ est nulle, donc est égale à un multiple de $n = 0 \times n$) ;

— symétrique : si x est congru à y modulo n (c'est-à-dire si la différence $x - y$ est un multiple de n), alors y est congru à x modulo n (car la différence $y - x$, qui est l'opposée de $x - y$, est aussi un multiple de n) ;

— transitive : si x est congru à y modulo n (différence $y - x$ multiple de n) et si y est congru à z modulo n (différence $z - y$ multiple de n), alors x est congru à z modulo n (car la différence $z - x$, qui est égale à $(z - y) + (y - x)$, est la somme de deux multiples de n, donc encore un multiple de n).

Reprenons l'exemple de la relation de congruence modulo 3 et cherchons les classes d'équivalence des éléments de $\mathbb{Z}$.
Dans la classe de 0, il y a tous les nombres entiers dont la différence avec 0 est un multiple de 3, c'est-à-dire tous les multiples de 3 :

$$\mathrm{Cl}(0) = \{ \ldots, -6, -3, 0, 3, 6, \ldots \}.$$

Dans la classe de 1, il y a tous les nombres entiers dont la différence avec 1 est un multiple de 3 :

$$\mathrm{Cl}(1) = \{ \ldots, -5, -2, 1, 4, 7, \ldots \}.$$

Dans la classe de 2, il y a tous les nombres entiers dont la différence avec 2 est un multiple de 3 :

$$\mathrm{Cl}(2) = \{ \ldots, -4, -1, 2, 5, 8, \ldots \}.$$

Tous les entiers ont été rangés dans ces trois classes. Il n'y a donc que trois classes d'équivalence : Cl(0), Cl(1), Cl(2) ; l'ensemble de ces trois classes est noté :

$\mathbb{Z}/3$ ou encore $\mathbb{Z}/3\mathbb{Z}$.

De façon générale, pour la relation de congruence modulo n, il y a n classes d'équivalence : celles de 0, de 1, de 2, ..., de $n - 1$. L'ensemble de ces classes d'équivalence est noté :

$\mathbb{Z}/n$ ou encore $\mathbb{Z}/n\mathbb{Z}$

Dans l'exemple de la relation de congruence modulo 3, si on fait la somme de :

— deux éléments quelconques de la classe de 0, on trouve un élément de la classe de 0 ;

— un élément de la classe de 0 avec un élément de la classe de 1, on trouve un élément de la classe de 1 ;

— un élément de la classe de 0 avec un élément de la classe de 2, on trouve un un élément de la classe de 2 ;

— deux éléments de la classe de 1 avec un élément de la classe de 2, on trouve un élément de la classe de 2 ;

— un élément de la classe de 1 avec un élément de la classe de 0, on trouve un élément de la classe de 1 ;

— un élément de la classe de 1 avec un élément de la classe de 2, on trouve un élément de la classe de 0 ;

— un élément de la classe de 2 avec un élément de la classe de 0, on trouve un élément de la classe de 2 ;

— un élément de la classe de 2 avec un élément de la classe de 1, on trouve un élément de la classe de 0 ;

— deux éléments de la classe de 2 avec un élément de la classe de 1, on trouve un élément de la classe de 1.

Ces propriétés nous permettent de définir la somme de deux quelconques de ces trois classes de la façon suivante :

$$\mathrm{Cl}(0) + \mathrm{Cl}(0) = \mathrm{Cl}(0)$$
$$\mathrm{Cl}(0) + \mathrm{Cl}(1) = \mathrm{Cl}(1) + \mathrm{Cl}(0) = \mathrm{Cl}(1)$$
$$\mathrm{Cl}(0) + \mathrm{Cl}(2) = \mathrm{Cl}(2) + \mathrm{Cl}(0) = \mathrm{Cl}(2)$$
$$\mathrm{Cl}(1) + \mathrm{Cl}(1) = \mathrm{Cl}(2)$$
$$\mathrm{Cl}(1) + \mathrm{Cl}(2) = \mathrm{Cl}(2) + \mathrm{Cl}(1) = \mathrm{Cl}(0)$$
$$\mathrm{Cl}(2) + \mathrm{Cl}(2) = \mathrm{Cl}(1)$$

Table d'addition de $\mathbb{Z}/3$

+	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

Cette addition de classes d'équivalence est une loi de composition interne dans l'ensemble $\mathbb{Z}/3$, dont la table est représentée ci-contre (en notant $\bar{0}$, $\bar{1}$, $\bar{2}$ les classes). En utilisant un procédé analogue avec le produit, on peut définir le produit de deux classes d'équivalence. Ce produit est une loi de composition interne dans $\mathbb{Z}/3$, dont la table est représentée ci-contre.

Table de multiplication de $\mathbb{Z}/3$

×	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Ce qu'on vient de faire dans le cadre de $\mathbb{Z}/3$ peut être fait pour un nombre entier n positif quelconque : on peut définir sur $\mathbb{Z}/n$ une addition et une multiplication. Ces deux lois de composition interne donnent à $\mathbb{Z}/n$ une structure d'anneau (et même une structure de corps si le nombre entier n est un nombre premier) ◆.

◆ Voir l'article « l'Algèbre », page

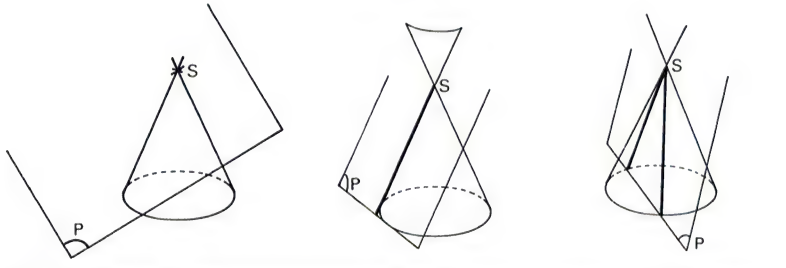
CONIQUES pages 326, 328.

Courbes intersection d'un cône de révolution et d'un plan. Cette intersection, si le plan passe par le sommet du cône, peut être :

● un point : le point S

● une génératrice (si ce plan est tangent au cône)

● deux génératrices



Si le plan ne passe pas par le sommet du cône, on appelle P ce plan. Par le sommet S du cône passe un plan et un seul parallèle à P ; soit Q ce plan. Si l'intersection de Q et du cône est un point, celle de P et du cône est une ellipse (un cercle dans le cas particulier où P est perpendiculaire à l'axe du cône). Si l'intersection de Q et du cône est une génératrice, celle de P et du cône est une parabole. Si l'intersection de Q et du cône est formée de deux droites, l'intersection de P et du cône est une hyperbole.

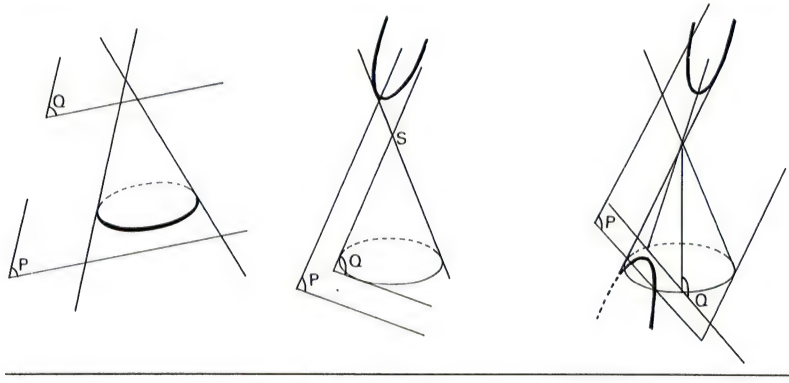
Toutes ces courbes ont une équation du second degré, de la forme :

$$ax^2 + by^2 + cxy + dx + ey + f = 0$$

Dans le cas où on obtient une droite (double) ou deux droites, on dit que la conique est décomposée ; dans les autres cas, on dit qu'elle est propre (parabole, hyperbole, ellipse).

Hyperbole et ellipse admettent un centre de symétrie et sont pour cela qualifiées de coniques à centre.

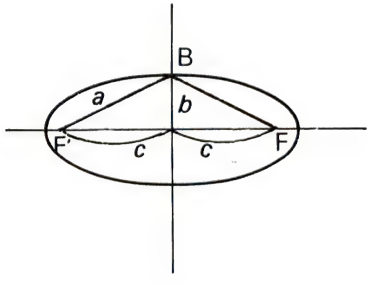
Ellipse **Parabole** **Hyperbole**

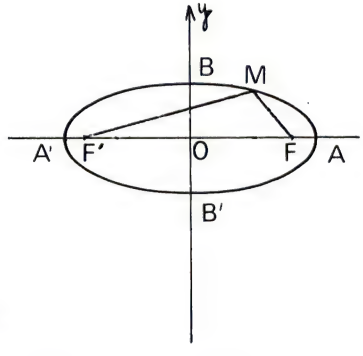


Ellipse (géométrie plane)

On appelle ellipse de foyers F et F′ (F et F′ étant deux points distincts du plan dont la distance est notée $2c$) l'ensemble des points M du plan tels que la somme

des distances de M à F et de M à F′ soit constante et égale à un nombre $2a$ (différent de 0 et supérieur à $2c$).
La droite FF′ et sa médiatrice sont axes de symétrie pour l'ellipse. Leur point d'intersection O est centre de symétrie.
Sur la droite passant par FF′, il y a 2 points de l'ellipse, A et A′ ; la distance de A à A′ est égale à $2a$. Sur la médiatrice de FF′, il y a aussi deux points de l'ellipse, B et B′. La distance de F à B = distance de F′ à B = a.

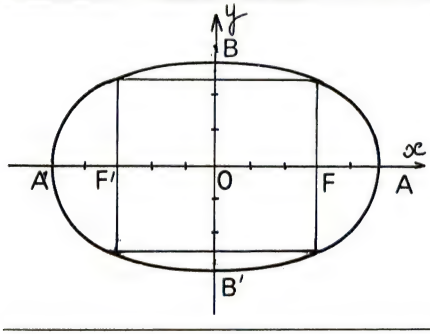


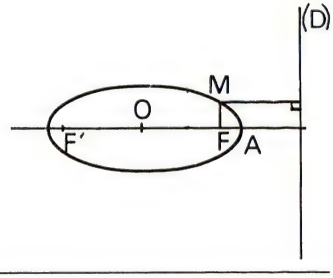


A, B, A′, B′ sont les sommets de l'ellipse.
AA′ est appelé grand axe de l'ellipse.
BB′ est appelé petit axe de l'ellipse.
$2a$ est la longueur du grand axe, $2b$ celle du petit, et $2c$ la distance de F à F′.
On a entre ces 3 nombres la relation $b^2 + c^2 = a^2$.
L'équation de l'ellipse dans des axes Ox, Oy (Ox = FF′, Oy = BB′) est :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

Il existe une droite D telle que l'ellipse soit aussi l'ensemble des points M dont le rapport des distances de M à F et de M à (D) soit constant (égal à un nombre e, appelé excentricité de l'ellipse, dont la valeur est $\frac{c}{a}$). Cette droite D est appelée directrice de l'ellipse associée au foyer F (en raison de la symétrie de la figure, il existe également une directrice associée au foyer F′ et qui a les mêmes propriétés que D). La distance de O à (D) est égale à $\frac{a^2}{c}$.





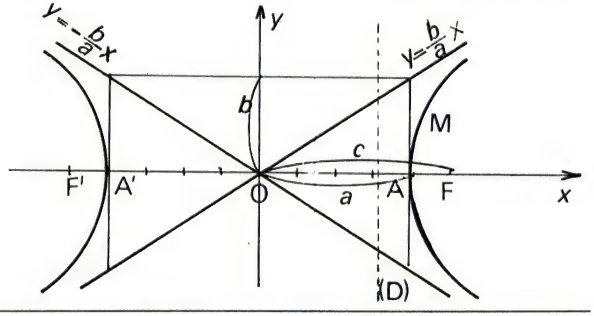
Hyperbole (géométrie plane)

Étant donné deux points distincts F et F′ dont la distance est $2c$ et un nombre, noté traditionnellement $2a$, on appelle hyperbole de foyers F et F′ l'ensemble des points M du plan dont la différence des distances de M à F et F′ est constante et égale à $2a$ (cela nécessite que $c > a$, sinon cet ensemble serait vide).
L'hyperbole est formée de deux branches :
— l'une pour laquelle la distance de M à F est supérieure à la distance de M à F′, où l'on a donc :

$$d(M, F) - d(M, F') = 2a$$

— l'autre pour laquelle c'est la distance de M à F′ qui est supérieure à celle de M à F : $d(M, F') - d(M, F) = 2a$.

Ces 2 branches sont symétriques par rapport à la médiatrice de FF′.

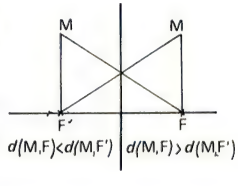


La droite FF′ et la médiatrice de FF′ sont axes de symétrie pour l'hyperbole. Le point O, milieu de FF′, est centre de symétrie. Il y a deux points de l'hyperbole sur la droite FF′ : A et A′ tels que $OA = OA' = a$; ces points sont appelés sommets de l'hyperbole. L'axe AA′ est appelé axe transverse. Il n'y a aucun point de l'hyperbole sur la médiatrice de FF′ ; cet axe est appelé axe non transverse.
Dans le repère Ox, Oy (Ox porté par FF′ et Oy par sa médiatrice), l'équation de l'hyperbole est :

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$$

où b est le nombre positif tel que $c^2 - a^2 = b^2$.

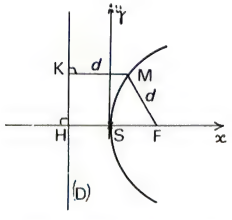
L'hyperbole admet 2 asymptotes ; ce sont les 2 droites d'équation $y = \frac{b}{a}x$ et $y = -\frac{b}{a}x$ dans le repère défini ci-dessus.



Il existe une droite D telle que l'hyperbole soit l'ensemble des points dont le rapport des distances à F et à (D) soit constant (ce nombre constant est appelé excentricité de l'hyperbole ; il est égal à $\frac{c}{a}$). Cette droite est la directrice de l'hyperbole associée au foyer F. Elle est parallèle à Oy et telle que la distance de O à (D) soit égale à $\frac{a^2}{c}$.

En raison de la symétrie, il y a aussi une directrice associée au foyer F′ et qui a les mêmes propriétés que D.

Parabole (géométrie plane)
Étant donné une droite (D) et un point F non situé sur D, on appelle parabole de foyer F et de directrice D l'ensemble des points M du plan équidistants de F et de D :



$d(M, F) = d(M, K)$
La droite passant par F perpendiculaire à (D) est axe de symétrie pour la figure. Sur cet axe, un seul point appartient à la parabole, le milieu S de HF. Ce point est appelé sommet de la parabole.

Dans un système d'axes Sx, Sy (Sx passant par F, et Sy paralllèle à D), l'équation de la parabole est $y = 2px$. La constante p est le *paramètre* de la parabole ; elle est égale à la distance de F à (D).

CONJONCTION : voir Et

CONJUGUÉ

Nombre complexe conjugué
Un nombre complexe s'écrit $x + iy$, où x et y sont deux nombres réels et i un nombre « imaginaire » dont le carré est égal à -1. On appelle nombre complexe conjugué du nombre $x + iy$ (ou simplement « conjugué de z ») le nombre complexe (unique) obtenu en gardant x tel quel et en changeant y de signe, c'est-à-dire le nombre $x - iy$. Si on note z le nombre complexe $x + iy$, on note $\overline{z}$ (z barre) le nombre complexe conjugué.

Propriétés du nombre complexe conjugué
z et z' étant deux nombres complexes.

1) Le conjugué de la somme $z + z'$ est égal à la somme des conjugués de z et de z' :
$$\overline{(z + z')} = \overline{z} + \overline{z'}.$$

2) Le conjugué du produit $z.z'$ est égal au produit des conjugués de $\overline{z}$ et $\overline{z'}$:
$$\overline{(z.z')} = \overline{z}\,.\,\overline{z'}.$$

3) Le conjugué du conjugué de z est égal à z :
$$\overline{\overline{z}} = z.$$

Les propriétés (1) et (2) prouvent que l'application de $\mathbb{C}$ dans $\mathbb{C}$ ($\mathbb{C}$ = ensemble des nombres complexes), qui à tout nombre complexe z fait correspondre son conjugué $\overline{z}$, est un endomorphisme de $\mathbb{C}$.

4) Si z est un nombre complexe non nul, il admet un inverse $\frac{1}{z}$ non nul et le conjugué de $\frac{1}{z}$ est l'inverse du conjugué de z :
$$\overline{\left(\frac{1}{z}\right)} = \frac{1}{\overline{z}}.$$

Matrice conjuguée

Si M est une matrice dont les coefficients sont des nombres complexes, on appelle conjuguée de M et on note $\overline{M}$ la matrice dont les coefficients sont les nombres complexes conjugués des coefficients de M :

$$M = \begin{pmatrix} 1+i & 3-2i \\ 4-i & 3i \end{pmatrix} \Longrightarrow \overline{M} = \begin{pmatrix} 1-i & 3+2i \\ 4+i & -3i \end{pmatrix}.$$

1) La matrice conjuguée d'une somme de deux matrices est égale à la somme des conjuguées des deux matrices :

$$\overline{M+N} = \overline{M} + \overline{N}.$$

2) La matrice conjuguée du produit de deux matrices est égale au produit des conjuguées des deux matrices :

$$\overline{M \times N} = \overline{M} \times \overline{N}.$$

3) Si λ est un nombre complexe quelconque, la matrice conjuguée de $\lambda \,.\, M$ est égale au produit par le conjugué de λ de la matrice conjuguée de M :

$$\overline{\lambda.M} = \overline{\lambda}.\overline{M}.$$

4) La matrice conjuguée de la matrice conjuguée $\overline{M}$ est la matrice M :

$$\overline{\overline{M}} = M.$$

5) Si M admet une matrice inverse M^{-1}, la matrice conjuguée de M^{-1} est l'inverse de la matrice conjuguée de M :

$$\overline{M^{-1}} = (\overline{M})^{-1}.$$

CONNEXE◆

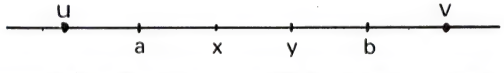
◆ Lire d'abord Intervalle.

Soit un intervalle quelconque $[a, b]$ de $\mathbb{R}$, ensemble des nombres réels. Il n'est pas possible de trouver des intervalles ouverts de $\mathbb{R}$, sans points communs, et dont la réunion soit égale à $[a, b]$.

Si on considère, par exemple, l'intervalle ouvert ◆ $]x, y[$ avec $a < x$ et $y < b$, un intervalle ouvert dont l'intersection avec $]x, y[$ soit vide est un intervalle dont origine et extrémité sont toutes les deux inférieures à x ou toutes les deux supérieures à y.

◆ Rappelons que, dans un intervalle ouvert $]x, y[$, x et y n'appartiennent pas à l'intervalle.

Si on prend $]u, x[$ avec $u \leqslant a$ et $]y, v[$ avec $v \geqslant b$, la réunion $]u, x[\cup]x, y[\cup]y, v[$ ne contient ni x ni y ; cette réunion n'est pas égale à $[a, b]$.



Un espace topologique ◆ E est connexe s'il n'est pas possible de trouver deux ouverts non vides de E dont l'intersection soit vide et dont la réunion soit égale à E.

◆ Voir Topologie.

Cette définition est équivalente à chacune des deux propriétés suivantes :

1) E n'est pas égal à la réunion de deux fermés disjoints.

2) Les seules parties de E qui soient à la fois ouvertes et fermées sont l'ensemble vide $\varnothing$ et l'ensemble E.

CONNEXION : voir Logique

CONSTANTE (Application)◆ pages 115, 125, 312.

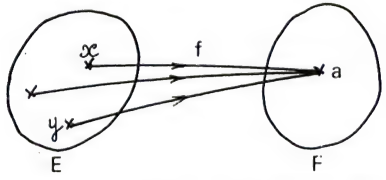
◆ Lire d'abord Application.

Une application f d'un ensemble E dans un ensemble F est dite constante si, pour tout couple (x, y) d'éléments de E, on a :

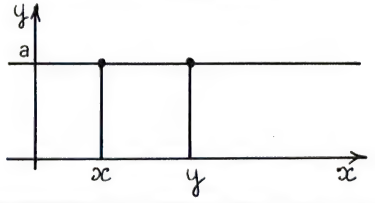
$f(x) = f(y)$.

Tous les éléments de E ont la même image par f : a ($a \in$ F) ; on dit souvent (par abus de langage) que l'application f est constante et égale à a.

Représentation sagittale



Représentation cartésienne



CONSTANTE (Fonction) : voir Fonction et la page 137.

CONTINU, CONTINUITÉ : voir l'article « l'Analyse », pages 118 à 124 et les pages 149, 311, 329, 369, 530, 539.

CONVERGENCE : voir l'article « l'Analyse », pages 109, 110, 113, et la page 539.

CONVEXE : page 404.

Une partie P du plan est convexe si, pour tout couple (M, N) de points de P, le segment MN est contenu dans P.

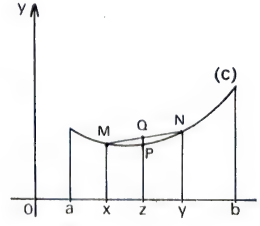
Fonction convexe

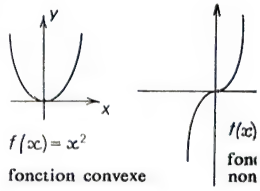
Soit une fonction définie sur un intervalle (a, b) de $\mathbb{R}$, ensemble des nombres réels, et soit (C) sa courbe représentative.

Si M est le point de (C) d'abscisse x et N le point de (C) d'abscisse y, on dit que la fonction est convexe si, pour tout nombre z compris entre x et y, le point de (C) d'abscisse z : P, est situé au-dessous du point Q de la droite MN qui a pour abscisse z.

$f(x) = x^2$: fonction convexe.

$f(x) = x^3$: fonction non convexe sur un intervalle de la forme $(-a, +a)$.





fonction convexe

non

COORDONNÉES : pages 247, 290, 291, 318, 329.

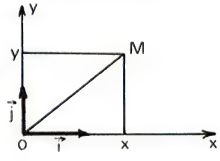
En géométrie plane

1) *Coordonnées cartésiennes*

Si $(O, \vec{i}, \vec{j})$ est un repère cartésien du plan, tout vecteur $\overrightarrow{OM}$ peut se mettre de façon unique sous la forme : $\overrightarrow{OM} = x\vec{i} + y\vec{j}$,

où x et y sont deux nombres réels ◆. Les nombres x et y sont appelés coordonnées cartésiennes du point M (x son abscisse, y son ordonnée). A tout couple (x, y) correspond un point et un seul ; à tout point M correspond un seul couple (x, y).

◆ Voir l'article « l'Algèbre linéaire », page 57.



2) *Coordonnées polaires*

Si $(\vec{O}, \vec{i}, \vec{j})$ est un repère cartésien orthonormé du plan, on peut repérer un point M du plan par :

— l'angle θ formé par $\overrightarrow{Ox}$ avec un axe porté par OM $(\overrightarrow{OX})$;
— la mesure algébrique de OM sur cet axe (ρ).

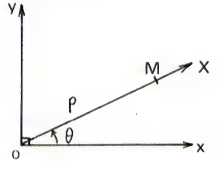
Ces deux nombres ρ et θ sont appelés coordonnées polaires de M.

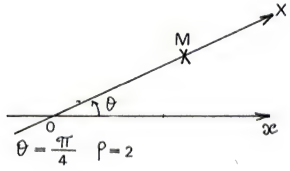
— A tout couple (ρ, θ) correspond un point M et un seul.
— A tout point M correspond une infinité ◆ de valeurs de ρ et θ :

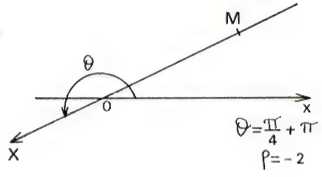
$$(\rho, \theta + 2k\pi) \text{ ou } (-\rho, \theta + (2k + 1)\pi),$$

avec k entier relatif quelconque.

◆ Chaque fois que l'angle θ varie de 2π, l'axe OX se retrouve dans la même position.







C'est le même point M qui correspond à ces deux systèmes de coordonnées polaires.

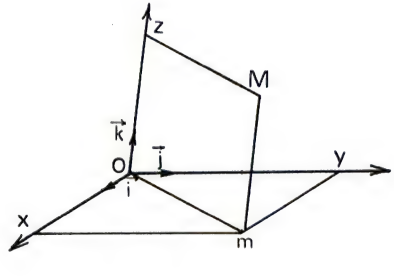
Les formules suivantes permettent le passage des coordonnées polaires aux coordonnées cartésiennes, et réciproquement :

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases} \quad \text{passage polaires} \longrightarrow \text{cartésiennes}$$

$$\begin{cases} \rho^2 = x^2 + y^2 \\ \cos \theta = \dfrac{x}{\rho} \\ \sin \theta = \dfrac{y}{\rho} \end{cases} \quad \text{passage cartésiennes} \longrightarrow \text{polaires}$$

En géométrie dans l'espace

1) *Coordonnées cartésiennes* (x, y, z)



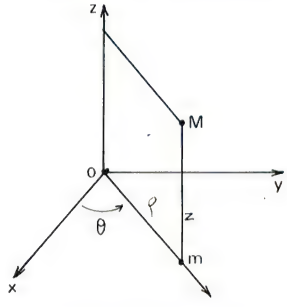
Si $(O, \vec{i}, \vec{j}, \vec{k})$ est un repère de l'espace, tout vecteur $\overrightarrow{OM}$ peut se mettre de façon unique sous la forme :

$$\overrightarrow{OM} = x\vec{i} + y\vec{j} + z\vec{k},$$

où x, y, z sont des nombres réels.

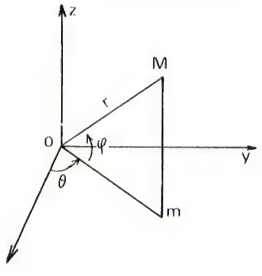
x, y, z sont appelés coordonnées cartésiennes de M (x = abscisse, y = ordonnée, z = cote).

2) *Coordonnées semi-polaires (ou cylindriques)*



Si $(O, \vec{i}, \vec{j}, \vec{k})$ est un repère orthonormé de l'espace, les coordonnées semi-polaires d'un point M sont les trois nombres ρ, θ, z définis de la façon suivante :
— ρ et θ sont les coordonnées polaires de la projection orthogonale m de M sur le plan xOy ;
— z est la cote de M (autrement dit, la mesure algébrique de mM).

3) *Coordonnées sphériques*



Si $(O, \vec{i}, \vec{j}, \vec{k})$ est un repère orthonormé de l'espace, les coordonnées sphériques d'un point M sont les trois nombres r, θ, φ ainsi définis :
— r est le nombre positif ou nul mesurant la distance de O à M ;
— θ est l'angle de $\overrightarrow{Ox}$ avec l'axe $\overrightarrow{Om}$ (m : projection orthogonale de M sur le plan xOy) ;
— φ est l'angle de $\overrightarrow{Om}$ avec $\overrightarrow{Ox}$.

4) *Coordonnées homogènes*

Soit M un point du plan (ou de l'espace) de coordonnées cartésiennes x, y (ou x, y, z dans l'espace). Soient X, Y, T (ou X, Y, Z, T dans l'espace) des nombres réels non tous nuls tels que :

$$x = \frac{X}{T}, \quad y = \frac{Y}{T} \quad \left(x = \frac{X}{T}, \quad y = \frac{Y}{T}, \quad z = \frac{Z}{T}\right).$$

Ces nombres sont appelés coordonnées homogènes du point M. Ces coordonnées ne sont pas définies de manière unique (en effet, si X, Y, T sont un système de coordonnées homogènes de M, k.X, k.Y, k.T en sont un autre système).
Un point M pour lequel T = 0 est appelé point à l'infini ◆. Tous les points M pour lesquels T = 0 constituent la droite de l'infini dans le cas de géométrie plane, le plan de l'infini dans le cas de géométrie dans l'espace. Deux droites parallèles se coupent au même point à l'infini ; deux plans parallèles se coupent suivant la même droite de l'infini.
L'espace ainsi muni de ces éléments à l'infini est appelé espace projectif (ou espace arguésien).

◆ Car, lorsque T tend vers 0, les valeurs de x, y (ou z) tendent vers l'infini.

COPLANAIRE

Des points, des droites, des vecteurs qui appartiennent à un même plan sont dits coplanaires.

CORPS : voir l'article « l'Algèbre », pages 53, 54 et les pages 60, 61, 102, 329, 371, 373, 374, 540.

CORRÉLATION : voir l'article « les Statistiques », pages 473 à 484 et les pages 176, 235.

COSINUS (Fonction) : voir Trigonométriques (fonctions)

COTE : voir Coordonnées

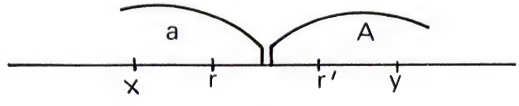
COUPLE pages 60, 61, 78, 207, 218, 232, 268, 387, 397.

Objet mathématique formé à partir de deux objets x et y, que l'on note (x, y). x s'appelle première composante du couple et y deuxième composante du couple. Deux couples (x, y) et (x', y') sont égaux si et seulement si $x = x'$ et $y = y'$. Donc, en général (si x et y ne sont pas égaux), les deux couples (x, y) et (y, x) sont distincts ◆. Parfois, pour insister sur cette importance de l'ordre des éléments x et y, on donne le nom de couple ordonné à un couple.

◆ Voir l article « les Ensembles », page où un exemple est donné.

COUPURE (de Dedekind)

On a effectué une coupure dans $\mathbb{Q}$ (ensemble des nombres rationnels) lorsqu'on a partagé $\mathbb{Q}$ en deux parties non vides, disjointes, a et A, appelées respectivement classe inférieure et classe supérieure, telles que :



— si r est élément de a, tout élément x de $\mathbb{Q}$ inférieur à r appartient à a ;

— si r' est élément de A, tout élément y de $\mathbb{Q}$ supérieur à r' appartient à A ;

— pour tout rationnel positif ε, on peut trouver au moins un x de a et un y de A tels que : $| y - x | < \varepsilon$.

Tout élément de A est supérieur à tout élément de a ; s'il existe un rationnel n'appartenant ni à a ni à A, il est supérieur à tout élément de a et inférieur à tout élément de A ; il ne peut exister deux rationnels non classés.
Il existe plusieurs types de coupures :

Premier type : coupure ouverte

Il existe un rationnel r qui n'appartient ni à a ni à A ; a est l'ensemble des rationnels strictement inférieurs à r ; A est l'ensemble des rationnels strictement supérieurs à r :

$$a = \{ x \mid x \in \mathbb{Q}, x < r \}$$

$$A = \{ x \mid x \in \mathbb{Q}, x > r \}.$$

Deuxième type : coupure fermée

— Tous les rationnels sont rangés soit dans l'une, soit dans l'autre des deux classes, et a admet un élément supérieur à tous les autres :

$$a = \{ x \mid x \in \mathbb{Q}, x \leqslant r \}$$
$$A = \{ x \mid x \in \mathbb{Q}, x > r \}.$$

— Tous les rationnels sont classés et A admet un élément inférieur à tous les autres :

$$a = \{ x \mid x \in \mathbb{Q}, x < r \}$$
$$A = \{ x \mid x \in \mathbb{Q}, x \geqslant r \}.$$

Troisième type : tous les rationnels sont classés ; a n'admet pas de plus grand élément, A n'admet pas de plus petit élément.

Exemple : a est l'ensemble des rationnels négatifs et des positifs dont le carré est inférieur à 2 ; A est l'ensemble des rationnels positifs dont le carré est supérieur à 2 :

$$a = \{ x \mid x \in \mathbb{Q}^- \} \cup \{ x \mid x \in \mathbb{Q}^+, x^2 < 2 \}$$
$$A = \{ x \mid x \in \mathbb{Q}^+, x^2 > 2 \}.$$

Dans ce cas, la coupure définit le nombre irrationnel $\sqrt{2}$; dans les autres cas, elle définit le rationnel r. Toute coupure dans $\mathbb{Q}$ définit un nombre réel (rationnel ou irrationnel).

COURBE CYCLOÏDALE : voir Cycloïde

COVARIANCE◆ page 435.

◆ Pour le lecteur non familiarisé avec les notions de statistiques, il est conseillé de lire l'article « les Statistiques », et notamment la corrélation.

Dans le cas d'une série statistique à deux variables X et Y, on peut déterminer les espérances mathématiques de chaque variable : E(X) et E(Y), ainsi que leur variance :

$$\frac{1}{n} \Sigma [X_i - E(X)]^2 \quad \text{et} \quad \frac{1}{n} \Sigma [Y_i - E(Y)]^2,$$

n étant le nombre d'éléments de la série.
La dispersion simultanée des deux variables est appelée moment produit ou covariance. Elle est égale à :

$$\frac{1}{n} \Sigma [X_i - E(X)] \; [Y_i - E(Y)].$$

Cette valeur intervient dans le calcul du coefficient de corrélation.
La covariance est nulle si les deux variables sont indépendantes.

Covariation

Ce terme désigne la liaison éventuelle existant entre les termes de deux séries chronologiques.

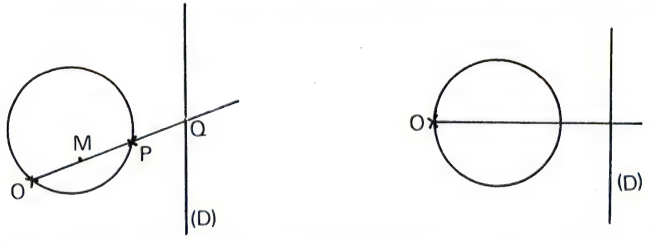
CROISÉ : voir Polygones

CROISSANTE (Fonction) : voir l'article « l'Analyse », page 112.

CUBE : voir Polyèdres

CUBIQUE (Géométrie plane)

Courbe plane dont l'équation $f(x, y) = 0$ est un polynôme du troisième degré.
Parmi les cubiques planes, citons les cubiques circulaires ; on peut les définir à partir d'un cercle et d'une droite de la façon suivante : si O est un point donné du cercle, une droite passant par O coupe le cercle en P et la droite en Q. Le point M tel que $\overrightarrow{OM} = \overrightarrow{PQ}$ décrit une cubique circulaire.
Si la droite (D) est parallèle à la tangente en O au cercle, la cubique admet pour axe de symétrie le diamètre du cercle qui passe par O ; on dit alors que c'est une cubique droite.



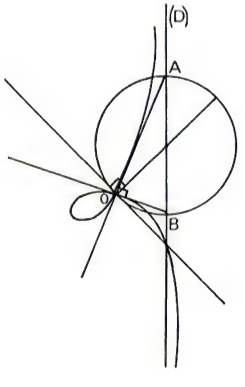
Si la droite (D) coupe le cercle en deux points A et B, la cubique admet O pour point double (car deux points M viennent se confondre avec O, dont les tangentes sont OA et OB : celui pour lequel P = Q = A et celui pour lequel P = Q = B).
Si la droite (D) est un diamètre du cercle, la courbe admet en O un point double à tangentes perpendiculaires : c'est une strophoïde.
Si la droite (D) est tangente au cercle (donc ne le coupe qu'en un seul point), la cubique admet un point de rebroussement en O : c'est une cissoïde.
Dans tous les cas, la droite (D) est asymptote à la cubique.

Équation de la cubique dans un repère d'origine O et d'axe Oy parallèle à (D) :

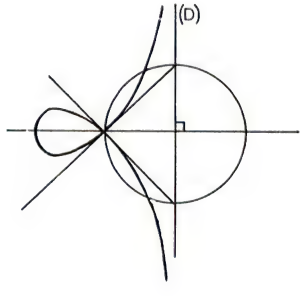
$$(x - a)(x^2 + y^2) + 2x(\alpha x + \beta y) = 0,$$

où $x = a$ est l'équation de (D), α et β les coordonnées du centre du cercle.

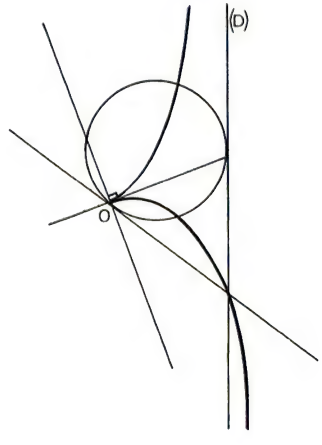
Strophoïde



Strophoïde droite



Cissoïde



CUMULATIF (Diagramme) : voir l'article « les Statistiques », page 462.

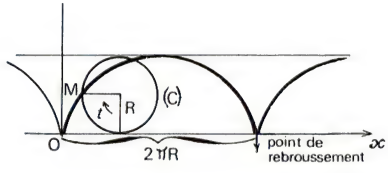
CYCLOÏDE page 196.

Lieu géométrique d'un point fixe M d'un cercle (C) qui roule, sans glisser, sur une droite Ox.

Équations paramétriques d'une cycloïde

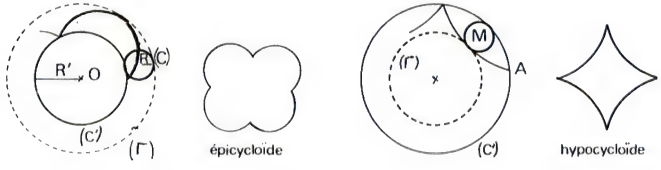
$$\begin{cases} x = R\,(t - \sin t) \\ y = R\,(1 - \cos t) \end{cases}$$

Une cycloïde admet une infinité de points de rebroussement situés sur la droite Ox ; la distance entre deux points de rebroussement consécutifs est $2\pi R$ (longueur du cercle).



Courbes cycloïdales

Lieux géométriques d'un point fixe M d'un cercle (C) qui roule sans glisser sur un cercle (C′).



1) Si le cercle (C), de rayon R, roule sans glisser sur un cercle (C′), de rayon R′, en étant extérieur au cercle (C′), la courbe décrite par M est une *épicycloïde*.

2) Si le cercle (C), de rayon R, roule sans glisser sur un cercle (C′), de rayon R′, en étant intérieur au cercle (C′), la courbe décrite par M est une *hypocycloïde* ◆.

◆ Remarquons que le cercle (C) reste aussi tangent à un cercle (Γ) extérieur à (C′) dans le cas d'une épicycloïde et intérieur à (C′) dans celui d'une hypocycloïde.

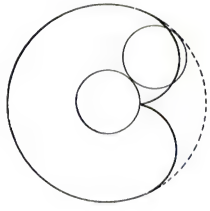
Équations paramétriques d'une courbe cycloïdale

Épicycloïde :
$$\begin{cases} x = (R + R')\cos t - R\cos\dfrac{R + R'}{R}\,t \\ y = (R + R')\sin t - R\sin\dfrac{R + R'}{R}\,t \end{cases}$$

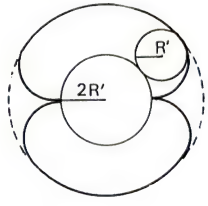
Hypocycloïde :
$$\begin{cases} x = (R - R')\cos t - R\cos\dfrac{R - R'}{R}\,t \\ y = (R - R')\sin t + R\sin\dfrac{R - R'}{R}\,t \end{cases}$$

Quelques courbes cycloïdales particulières

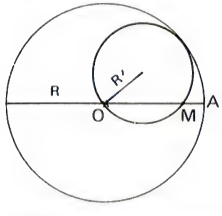
Cardioïde
$R = R'$
Épicycloïde à 1 rebroussement



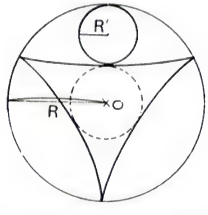
Néphroïde
$R = 2R'$
Épicycloïde à
2 rebroussements



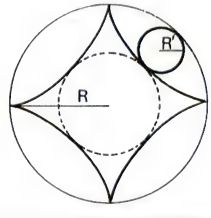
Droite
$R = 2R'$
Hypocycloïde à
2 rebroussements



$R = 3R'$
Hypocycloïde
à 3 rebroussements



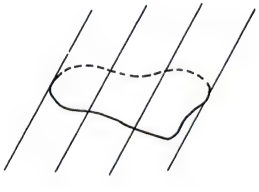
Astroïde
$R = 4R'$
Hypocycloïde
à 4 rebroussements



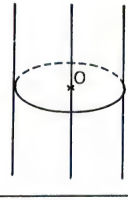
CYLINDRE, CYLINDRE DE RÉVOLUTION

Ensemble des droites parallèles à une droite (D) et rencontrant une courbe (C). Les droites sont appelées les génératrices du cylindre ; la courbe (C) est appelée directrice du cylindre.
Si la courbe (C) est un cercle et si les génératrices sont parallèles à l'axe du cercle, le cylindre est appelé cylindre de révolution.

Cylindre



Cylindre de révolution



D

DÉCAGONE : voir Polygones

DÉCIMAL (Nombre) pages 371 et 372, 444.

$\frac{3}{5}$ est un nombre décimal, car, si on le multiplie par 10, on obtient 6, qui est un nombre entier ; $\frac{17}{50}$ est aussi un nombre décimal, car, si on le multiplie par 10, le résultat obtenu est $\frac{17}{5}$, donc non entier ; mais, si on le multiplie par 100, le nombre obtenu est 34 qui, lui, est entier.

$\frac{2}{3}$ n'est pas un nombre décimal, car on peut le multiplier par 10, $10^2 = 1\,000$, $10^3 = 1\,000$, etc., par n'importe quelle puissance de 10 sans jamais obtenir un résultat entier.

Un nombre décimal x est donc un nombre rationnel tel qu'on puisse trouver une puissance de 10 (10, 100, 1 000, 10 000, ...) dont le produit par x soit un nombre entier.

Si on écrit un nombre décimal sous forme de fraction irréductible $\frac{p}{q}$, q est le produit uniquement de termes égaux à 2 ou à 5 ; quant à son développement décimal, il est fini. Par exemple, $\frac{31}{25}$ est un nombre décimal qui peut aussi être écrit 1,24.

DÉCROISSANCE : voir l'article « l'Analyse », pages 112 et 122.

DENSE ◆ pages 370, 371, 372, 373, 374.

◆ Lire Topologie, Adhérence.

Une partie A d'un espace topologique E est dense dans E si son adhérence est égale à E. Donc :

— l'extérieur de A est vide ;

— tout ensemble ouvert non vide de E a une intersection non vide avec A (ces deux propriétés sont, d'ailleurs, équivalentes à la définition).

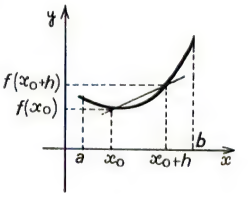
Exemple ◆

◆ Lire d'abord l'article « les Nombres ».

Si l'ensemble E est l'ensemble des nombres réels $\mathbb{R}$ et si A est l'ensemble $\mathbb{Q}$ des nombres rationnels, n'importe quel nombre réel x est adhérent à A (car l'intersection d'un voisinage quelconque de x avec $\mathbb{Q}$ n'est pas vide). L'adhérence de $\mathbb{Q}$ est donc $\mathbb{R}$ tout entier. $\mathbb{Q}$ est dense dans $\mathbb{R}$.

DÉRIVÉE ◆ pages 61, 80, 125 à 134, 143 à 147, 150, 157, 182, 245, 291, 292, 304, 311, 313, 319.

◆ Voir l'article « l'Analyse », pages 125 et suivantes.



Nombre dérivé d'une fonction pour une valeur x_0 de la variable

Si f est une fonction définie sur un intervalle (a, b) de $\mathbb{R}$ (ensemble des nombres réels), si x_0 est un point de l'intervalle (a, b) et si $x_0 + h$ en est un autre, on appelle nombre dérivé de la fonction f pour la valeur x_0 la limite (si elle existe et si elle est finie), lorsque h tend vers zéro, du rapport :

$$\frac{f(x_0 + h) - f(x_0)}{h}.$$

Ce nombre est noté $f'(x_0)$. On dit alors que f est dérivable en x_0.

Exemple

La fonction f définie sur $\mathbb{R}$ (ensemble de nombres réels) par $f(x) = x^2$ admet un nombre dérivé pour $x_0 = 3$; en effet, le rapport :

$$\frac{f(3 + h) - f(3)}{h} = \frac{(3 + h)^2 - (3)^2}{h} = \frac{9 + 6h + h^2 - 9}{h} = \frac{6h + h^2}{h} = 6 + h$$

admet pour limite 6 quand h tend vers 0. Le nombre dérivé de f pour la valeur $x_0 = 3$ est donc 6.

Fonction dérivée

Si la fonction f est dérivable en chaque point x_0 de (a, b), on dit qu'elle est dérivable sur (a, b) et on appelle fonction dérivée la fonction f' qui à chaque valeur x de (a, b) fait correspondre la valeur $f'(x)$ du nombre dérivé de f au point x.

Exemple

La fonction f de l'exemple précédent est dérivable pour n'importe quelle valeur x réelle, car le rapport :

$$\frac{f(x + h) - f(x)}{h} = \frac{(x + h)^2 - x^2}{h} = \frac{x^2 + 2hx + h^2 - x^2}{h} = \frac{2hx + h^2}{h} = 2x + h$$

admet pour limite $2x$ quand h tend vers 0.
La fonction dérivée de la fonction f est la fonction f' qui à x fait correspondre $2x$.

Interprétation géométrique de la dérivée

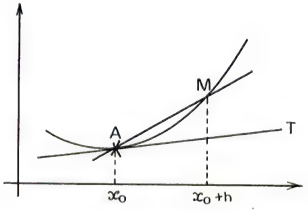
Si (C) est la courbe représentative de la fonction f, A le point d'abscisse x_0 de la courbe (C), le rapport

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

est la valeur du coefficient directeur de la droite AM (M étant le point d'abscisse $x_0 + h$ de la courbe (C)). Lorsque h tend vers 0, le point M tend vers le point A et, si la fonction f est dérivable pour x_0, le rapport

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

admet une limite finie ; donc la droite AM a une position limite AT qui est la tangente en A à la courbe (C) ; le coefficient directeur de la droite AT est la valeur de la dérivée au point x_0.



Exemple
La fonction $f(x) = x^2$ est dérivable pour $x_0 = 3$; son nombre dérivé est 6. La courbe représentative de cette fonction admet au point A d'abscisse 3 une tangente dont le coefficient directeur est 6.

Dérivées successives
Soit f une fonction dérivable sur l'intervalle (a, b) et f' sa fonction dérivée (cette fonction f' est quelquefois qualifiée de dérivée première). Si la fonction f' est elle-même dérivable sur (a, b), sa fonction dérivée est appelée fonction dérivée seconde de f et notée f'' ; si f'' est dérivable sur (a, b), sa fonction dérivée est appelée fonction dérivée troisième de f, etc.

Exemple

Si $f(x)$	$= x^4$	$f'''(x)$	$= 24x$
$f'(x)$	$= 4x^3$	$f''''(x)$	$= 24$
$f''(x)$	$= 12x^2$	$f^{(5)}(x)$	$= 0$, ainsi que toutes les dérivées suivantes.

DÉTERMINANT page 522.

Déterminant d'une matrice carrée (à coefficients réels)

— Si la matrice carrée M est d'ordre 2, soit $M = \begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$, le déterminant de M est le nombre réel, noté $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$, égal à $ab' - ba'$

$$\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - ba'.$$

— Si la matrice carrée M est d'ordre 3, soit $M = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$, le déterminant de M est le nombre réel, noté $\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix}$, égal à

$$a \times \begin{vmatrix} b' & c' \\ b'' & c'' \end{vmatrix} - b \times \begin{vmatrix} a' & c' \\ a'' & c'' \end{vmatrix} + c \begin{vmatrix} a' & b' \\ a'' & b'' \end{vmatrix}$$

c'est-à-dire à :

$$a(b'c'' - b''c') - b(a'c'' - a''c') + c(a'b'' - a''b')$$

ou encore :

$$ab'c'' + a'b''c + a''bc' - ac'b'' - a'bc'' - a''b'c.$$

Pour obtenir la valeur de ce déterminant, on a :
1. multiplié chaque élément de la première ligne par le déterminant d'ordre 2 obtenu en supprimant la ligne et la colonne de l'élément considéré ;
2. affecté le produit contenant a du signe +, le produit contenant b du signe —, le produit contenant c du signe + ;
3. fait la somme de ces 3 nombres.

On aurait pu opérer de la même façon à partir de la première colonne au lieu de la première ligne, ou même à partir d'une ligne ou d'une colonne quelconques, à condition d'affecter chaque élément de cette ligne ou de cette colonne d'un signe : + si la somme $i + j$ du numéro de la ligne et du numéro de la colonne est paire, — si cette somme est impaire.

Ainsi, la valeur du déterminant, calculée à partir de la deuxième colonne est la suivante :

$$-b\begin{vmatrix} a' & c' \\ a'' & c'' \end{vmatrix} + b'\begin{vmatrix} a & c \\ a'' & c'' \end{vmatrix} - b''\begin{vmatrix} a & c \\ a' & c' \end{vmatrix}$$
$$= -b(a'c'' - a''c') + b'(ac'' - a''c) - b''(ac' - a'c)$$
$$= ab'c'' + a'b''c + a''bc' - ac'b'' - a'bc'' - a''b'c$$

on constate que c'est bien la même.
b a été affecté du signe —, car il est dans la première ligne et la deuxième colonne ($1 + 2 = 3$, impair), b'' aussi, car il est dans la troisième ligne et la deuxième colonne ($3 + 2 = 5$, impair), tandis que b' est affecté du signe +, car il est dans la deuxième ligne et la deuxième colonne ($2 + 2 = 4$, pair).
— Si la matrice est d'ordre 4 (ou d'un ordre quelconque n supérieur à 3), on procède de la même façon. Les déterminants par lesquels on multiplie les éléments d'une ligne ou d'une colonne sont d'ordre 3 (ou d'ordre $n - 1$).

Déterminant d'un système d'équations linéaires

Le déterminant d'un système de deux équations linéaires à deux inconnues, de la forme : $\begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$, est le nombre : $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}$ $(= ab' - ba')$; celui d'un système de trois équations à trois inconnues, $\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \end{cases}$, est le nombre : $\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix}$.

Propriétés des déterminants

— Si tous les éléments d'une ligne, ou d'une colonne, sont nuls, le déterminant est nul.
— Si deux lignes, ou deux colonnes, ont leurs éléments correspondants égaux, le déterminant est nul.
— Si on multiplie tous les éléments d'une ligne, ou d'une colonne, par un même nombre, le déterminant est multiplié par ce nombre.
— Si on échange deux lignes, ou deux colonnes, la valeur du déterminant est changée de signe.
— On ne change pas la valeur du déterminant si on échange le rôle des lignes et des colonnes $\left(\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 5 \\ -1 & 2 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 1 & 2 \\ 3 & 5 & 4 \end{vmatrix} \right)$.

— On peut faire la somme ou le produit de deux déterminants de même ordre exactement de la même façon que l'on fait la somme ou le produit de deux matrices. À noter toutefois que le produit de deux déterminants peut se faire non seulement ligne par colonne (comme le produit de deux matrices), mais aussi ligne par ligne (ou colonne par colonne).

— Ainsi : $\begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} \times \begin{vmatrix} 4 & 0 \\ 5 & -2 \end{vmatrix}$

$$= \begin{vmatrix} 1 \times 4 + 2 \times 5 & 1 \times 0 + 2 \times (-2) \\ 3 \times 4 + (-1) \times 5 & 3 \times 0 + (-1) \times (-2) \end{vmatrix}$$

$$= \begin{vmatrix} 14 & -4 \\ 7 & 2 \end{vmatrix} = 28 + 28 = 56$$

cela est le produit fait ligne par colonne.

$$\begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} \times \begin{vmatrix} 4 & 0 \\ 5 & -2 \end{vmatrix}$$

$$= \begin{vmatrix} 1 \times 4 + 2 \times 0 & 1 \times 5 + 2 \times (-2) \\ 3 \times 4 + (-1) \times 0 & 3 \times 5 + (-1)(-2) \end{vmatrix}$$

$$= \begin{vmatrix} 4 & 1 \\ 12 & 17 \end{vmatrix} = 68 - 12 = 56$$

cela est le produit fait ligne par ligne.

Il est évident que, dans ce cas, il était plus rapide de calculer chacun des deux déterminants et d'en faire le produit :

$$(1 \times (-1) - 2 \times 3) \times (4 \times (-2) - 5 \times 0) = (-7) \times (-8) = 56.$$

— À noter aussi que la multiplication des déterminants est une opération commutative, contrairement à celle des matrices.

Par exemple, si $M = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$ et $N = \begin{pmatrix} 4 & 0 \\ 5 & -2 \end{pmatrix}$, le produit $M \times N$ des deux matrices est la matrice :

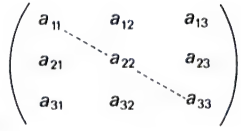
$$M \times N = \begin{pmatrix} 14 & -4 \\ 7 & 2 \end{pmatrix}$$ dont le déterminant vaut 56,

tandis que le produit $N \times M$ est la matrice :

$$N \times M = \begin{pmatrix} 4 & 8 \\ -1 & 12 \end{pmatrix}$$ dont le déterminant vaut 56.

DIAGONALE PRINCIPALE page 451.

On appelle diagonale principale d'un tableau carré l'ensemble des éléments dont l'indice de ligne et l'indice de colonne sont égaux.



◆ Voir aussi le début de l'article « les Ensembles ».

DIAGRAMME pages 260, 266, 268, 271, 314.

Schéma utilisé comme support du raisonnement ◆.

● *Diagramme de Venn ou d'Euler :* diagramme utilisé pour représenter des ensembles. Les éléments de ces ensembles sont figurés par des points entourés d'une courbe fermée (sans point double).

Il arrive qu'on ne représente plus les points (si les éléments sont en trop grand nombre, par exemple) et un ensemble est alors représenté par la partie du plan intérieure à la courbe.

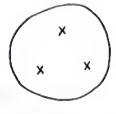


Diagramme de Venn ou d'Euler

● *Diagramme de Caroll.* Un ensemble E est représenté par la portion du plan intérieure à un carré et un sous-ensemble de E est obtenu en partageant ce carré à l'aide de droites parallèles aux côtés du carré.

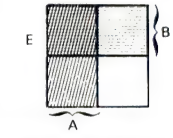


Diagramme de Caroll

● *Diagrammes d'une relation* $\mathcal{R}$ d'un ensemble E vers un ensemble F.

Diagramme cartésien. Les éléments de E sont représentés par des points d'une droite, ceux de F par des points d'une droite perpendiculaire. Si l'élément x de E est lié à y de F par la relation $\mathcal{R}$, on place une croix au point d'abscisse x et d'ordonnée y. La relation représentée ci-contre est celle qui est formulée par « a pour carré ». Exemple : l'élément 2 de E a pour carré l'élément 4 de F. L'écriture de cette représentation est celle qui est donnée sous le dessin.

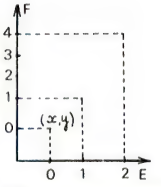
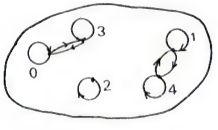


Diagramme cartésien de la relation « a pour carré » de E = { 0, 1, 2 } vers F = { 0, 1, 2, 3, 4 }.

Diagramme sagittal. Les ensembles E et F sont représentés par un diagramme de Venn ; chaque fois que l'élément x de E est lié par la relation $\mathcal{R}$ à l'élément y de F, on joint x à y par une flèche allant de x vers y.

Si les deux ensembles E et F sont égaux, on peut soit représenter E en deux exemplaires, soit en un seul (certaines propriétés des relations apparaissant cependant plus clairement si on représente E en un seul exemplaire).



Par exemple, soit la relation « a le même reste dans la division par 3 que » définie de l'ensemble E = { 0, 1, 2, 3, 4 } dans lui-même.

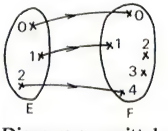
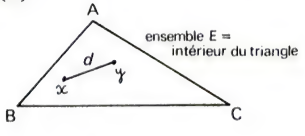


Diagramme sagittal de la même relation

DIAMÈTRE

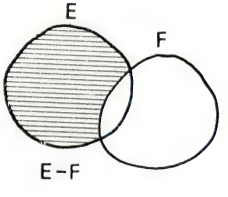
Si E est un ensemble muni d'une distance, notée d (on note alors $d(x, y)$ la distance de deux points x et y de E), on appelle diamètre de E la borne supérieure de l'ensemble constitué par tous les nombres $d(x\,y)$ que l'on peut obtenir. Ce nombre est noté δ(E).



Le diamètre de E est la distance BC : c'est la plus grande distance qui existe entre deux points de ce triangle.

DIFFÉRENCE page 223.

◆ On lit « E moins F ».



Différence de deux ensembles

La différence des deux ensembles E et F est l'ensemble, noté E — F ◆, des éléments de E qui n'appartiennent pas à F :

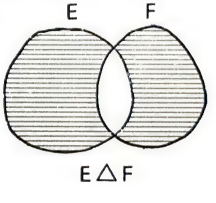
$$E - F = \{ x \mid x \in E \text{ et } x \notin F \}$$

Par exemple, si $E = \{ 1, 2, 3, 4 \}$ et $F = \{ 1, 2, 5 \}$, $E - F = \{ 3, 4 \}$

La différence E — F est aussi le complémentaire dans E de l'intersection E ∩ F :

$$E - F = \complement_E(E \cap F)$$

◆ On lit « E delta F ».



Différence symétrique de deux ensembles

La différence symétrique des deux ensembles E et F est l'ensemble, noté E Δ F ◆, des éléments qui appartiennent à un et un seul des deux ensembles ; c'est donc l'ensemble des éléments qui appartiennent à E sans appartenir à F ou qui appartiennent à F sans appartenir à E :

$$E \Delta F = \{ x \mid x \in E \text{ et } x \notin F \} \cup \{ x \mid x \in F \text{ et } x \notin E \}$$

C'est aussi la réunion de E — F avec F — E :

$$E \Delta F = (E - F) \cup (F - E)$$

Ou encore la différence de E ∪ F et de E ∩ F :

$$E \Delta F = (E \cup F) - (E \cap F)$$

Exemple

$$E = \{ 1, 2, 3, 4, 5 \}$$

$$F = \{ 2, 4, 6 \}$$

$$E \Delta F = \{ 1, 3, 5, 6 \}$$

$$\left.\begin{array}{l} E - F = \{ 1, 3, 5 \} \\ F - E = \{ 6 \} \end{array}\right\} (E - F) \cup (F - E) = \{ 1, 3, 5, 6 \}$$

$$\left.\begin{array}{l} E \cup F = \{ 1, 2, 3, 4, 5, 6 \} \\ E \cap F = \{ 2, 4 \} \end{array}\right\} (E \cup F) - (E \cap F) = \{ 1, 3, 5, 6 \}$$

La différence symétrique de E et de F est égale à la différence symétrique de F et de E :

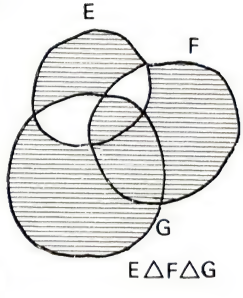
$$E \Delta F = F \Delta E$$

Si on fait intervenir un troisième ensemble G, on peut :
— soit calculer (E Δ F), puis (E Δ F) Δ G ;
— soit calculer (F Δ G), puis E Δ (F Δ G).

On trouve le même résultat :

$$(E \Delta F) \Delta G = E \Delta (F \Delta G)$$

On supprime les parenthèses et on écrit : E Δ F Δ G.



Quel que soit l'ensemble E, la différence symétrique de E et de l'ensemble vide ∅ est l'ensemble E lui-même :

$$E \Delta \varnothing = E$$

Et la différence symétrique de E et de lui-même est E :

$$E \Delta E = \varnothing$$

Différence de deux entiers naturels

a et *b* étant deux entiers naturels, s'il existe un entier naturel *c* tel que *b* soit égal à la somme de *a* et de *c*, cet entier *c* est unique ; on l'appelle différence de *b* et de *a* et on le note $b - a$:

$$a \in \mathbb{N}, \quad b \in \mathbb{N}, \quad c \in \mathbb{N} \qquad b = a + c \Longrightarrow c = b - a$$

Différence de deux éléments ◆

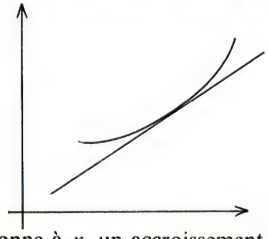
◆ Ce passage ne peut être compris que si le lecteur a lu l'article « l'Algèbre ».

E étant un ensemble muni d'une loi de composition interne (notée +) (commutative, associative et possédant un élément neutre noté O), si l'élément *a* de E admet un opposé $-a$, l'équation $a + x = b$ (où *b* est un élément de E) admet une solution dans E et une seule : $x = b + (-a)$, que l'on note $b - a$. Cet élément de E, $b - a$, est appelé différence de *b* et de *a* (cette différence n'existe donc que si *a* admet un opposé). Si E a une structure de groupe pour la loi +, alors, pour tout *a* et tout *b*, la différence $b - a$ existe.

DIFFÉRENTIELLE, DIFFÉRENTIABLE ◆ pages 319, 327.

◆ Lire d'abord Dérivée et Intervalle.

Soit f une fonction réelle de la variable réelle x, définie sur un intervalle I de $\mathbb{R}$ (ensemble des nombres réels), dérivable sur I, et soit x_0 un point de I.



La fonction f étant dérivable, sa courbe représentative (C) admet une tangente en chacun de ses points. Soit MT la tangente à (C) au point M d'abscisse x_0 ; l'équation de cette tangente est :

$$y = x.f'(x_0) + y_0 - x_0.f'(x_0).$$

Notons $y = g(x)$ cette équation.

Si on donne à x_0 un accroissement h (dans le cas du dessin, h est positif, mais il pourrait être négatif), l'accroissement correspondant de la fonction f est :

$$f(x_0 + h) - f(x_0) \qquad \text{(noté } \Delta y\text{)}$$

et l'accroissement correspondant de la fonction g est :

$$g(x_0 + h) - g(x_0) = h.f'(x_0) \qquad \text{(noté } \mathrm{d}y\text{)}.$$

On démontre que la différence entre Δy et $\mathrm{d}y$ tend vers 0 lorsque h tend vers 0. On dit que la fonction f est différentiable au point x_0 et que la différentielle de f au point x_0 est la fonction, notée $\mathrm{d}_{x_0}f$, qui à chaque valeur de h fait correspondre $h \times f'(x_0)$:

$$\mathrm{d}_{x_0}f : h \longmapsto h \times f'(x_0).$$

Si la fonction f est différentiable pour n'importe quelle valeur x de l'intervalle I, on dit qu'elle est différentiable sur I et que sa différentielle sur I est la fonction $\mathrm{d}f$ qui à chaque nombre réel h fait correspondre le nombre réel $h \times f'(x)$ ◆ :

$$\mathrm{d}f : h \longmapsto h \times f'(x).$$

◆ $\mathrm{d}f$ est donc une fonction de deux variables : h et x.

Le nombre $h.f'(x)$ est noté $\mathrm{d}y$. Le nombre h est noté $\mathrm{d}x$ ◆ et on a donc la relation

$$\mathrm{d}y = f'(x).dx.$$

◆ Cette relation $\mathrm{d}x$ est bien compatible avec le fait que, dans le cas de l'application identique ($x \longmapsto x$), l'équation $y = f(x)$ s'écrit $y = x$, donc $\mathrm{d}y = \mathrm{d}x$; et la relation $\mathrm{d}y = f'(x).\mathrm{d}x$ s'écrit bien aussi $\mathrm{d}y = \mathrm{d}x$, puisque la dérivée de la fonction identique est égale à 1.

Cette écriture peut s'écrire aussi :

$$f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} ;$$

cette écriture de la dérivée s'appelle notation différentielle de la dérivée.

DIMENSION : voir Algèbre linéaire

DISCONTINU : voir l'article « l'Analyse », pages 118, 119 et la page 182.

DISCRÈTE (Topologie) ◆

◆ Mot du vocabulaire de Topologie. Lire ce mot.

Topologie pour laquelle toute partie d'un ensemble E est un ouvert de E.

DISCRIMINANT : voir Équations (du second et du troisième degré) et les pages 286 à 288.

DISJOINTS

Deux ensembles E et F sont disjoints si leur intersection $E \cap F$ est vide (s'ils n'ont aucun élément commun).

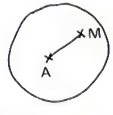
Ensembles disjoints

DISJONCTION : voir Ou

DISQUE

Dans le plan euclidien (c'est-à-dire plan où l'on a défini une distance ◆), on appelle *disque fermé* l'ensemble des points M du plan, dont la distance à un point A est *inférieure* ou *égale* à un nombre positif R. Le point A est appelé centre du disque et le nombre R, rayon du disque.
Si on ne considère que les points dont la distance à A est strictement inférieure à R, on dira *disque ouvert*.

◆ Lire ce mot.



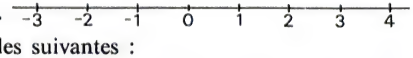
Distance de A à M $\leqslant$ R

DISTANCE (espace métrique)

pages 94 et 95, 243, 330, 345, 346, 454, 529, 531, 532.

Sur l'ensemble des nombres réels

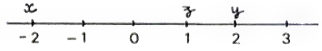
Associons à deux nombres réels quelconques x et y la différence entre le plus grand des deux et le plus petit. Cette différence est un nombre réel positif (qui n'est autre que la valeur absolue de la différence $x - y$).
On appellera ce nombre la distance de x à y et on notera $d(x, y)$. Par exemple :
$d(2, 3) = 3 - 2 = 1$;
$d(1, -2) = 1 - (-2) = 1 + 2 = 3$.



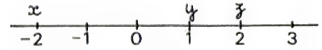
Les propriétés de cette distance sont les suivantes :
— la distance de x à y et la distance de y à x sont égales : $d(x, y) = d(y, x)$;
— la distance de x à y est nulle si $x = y$ et seulement dans ce cas ; si x est différent de y, alors la distance est strictement positive ;
— si z est un troisième nombre réel, la distance de x à y est inférieure ou égale à la somme des distances de x à z et de z à y : $d(x, y) \leqslant d(x, z) + d(z, y)$.

Exemples

$$\left.\begin{aligned} d(x, y) &= 1 - (-2) = 3 \\ d(x, z) &= 2 - (-2) = 4 \\ d(z, y) &= 2 - 1 \quad\;\; = 1 \end{aligned}\right\} \quad 3 < 4 + 1$$



$$\left.\begin{aligned} d(x, y) &= 4 \\ d(x, z) &= 3 \\ d(z, y) &= 1 \end{aligned}\right\} \quad 4 = 3 + 1$$



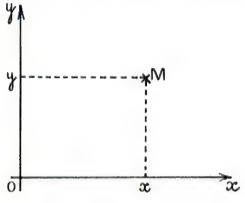
On dit avoir muni $\mathbb{R}$ d'une distance d. Dans le cas de cet exemple, la distance d est l'application qui à deux nombres réels x et y fait correspondre le nombre $d(x, y) = |x - y|$ (valeur absolue de $x - y$).

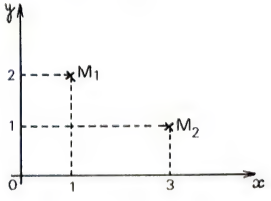
Sur un ensemble quelconque

De façon générale, étant donné un ensemble E, on dit qu'on a muni E d'une distance d si on a défini une application d qui à deux éléments quelconques x et y

de E fait correspondre un nombre réel positif ou nul $d(x, y)$ satisfaisant aux conditions :
1) $d(x, y) = d(y, x)$, quels que soient x et y.
2) $d(x, y) = 0$, si et seulement si $x = y$.
3) $d(x, y) \leqslant d(x, z) + d(z, y)$, quels que soient x, y, z.
On a vu ci-dessus qu'on pouvait effectivement munir $\mathbb{R}$ d'une distance. Étant donné un ensemble E, on peut munir E d'une distance de différentes façons ; nous allons le montrer dans le cas de $\mathbb{R}^2$ et des trois distances usuelles que l'on peut définir sur cet ensemble.
$\mathbb{R}^2$ ◆ est un ensemble de couples, un point de $\mathbb{R}^2$ est donc un couple (x, y). Pour plus de commodité, on représente le point (x, y) de $\mathbb{R}^2$ par le point du plan dont les coordonnées sont x et y (dans un certain repère donné).

◆ $\mathbb{R}^2$ est l'ensemble des couples (x, y) où x et y appartiennent tous les deux à $\mathbb{R}$.





La première distance d_1 est celle qui au point M_1 de coordonnées (x_1, y_1) et au point M_2 de coordonnées (x_2, y_2) fait correspondre le nombre :

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Cette première distance est appelée distance euclidienne (c'est celle qu'on utilise couramment).

Exemple : si M_1 a pour coordonnées (1, 2) et si M_2 a pour coordonnées (3, 1)' la distance de M_1 à M_2 est :

$$d_1(M_1, M_2) = \sqrt{(3 - 1)^2 + (1 - 2)^2} = \sqrt{4 + 1} = \sqrt{5}.$$

La deuxième distance d_2 est celle qui aux deux points M_1 et M_2 fait correspondre le nombre :

$$| x_1 - x_2 | + | y_1 - y_2 |.$$

Avec le même exemple :

$$d_2(M_1, M_2) = | 3 - 1 | + | 2 - 1 | = 2 + 1 = 3.$$

La troisième distance d_3 est celle qui aux deux points M_1 et M_2 fait correspondre le nombre : sup $(| x_1 - x_2 |, | y_1 - y_2 |)$, c'est-à-dire le plus grand des deux nombres $| x_1 - x_2 |$ et $| y_1 - y_2 |$.
Avec le même exemple :

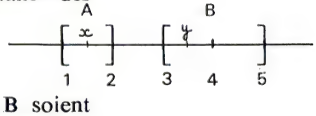
$$d_3(M_1, M_2) = \sup(| 3 - 1 |, | 2 - 1 |) = \sup(2, 1) = 2.$$

Ces trois applications sont effectivement des distances, car elles vérifient bien les propriétés (1), (2) et (3).
Un ensemble E sur lequel a été définie une distance est appelé un espace métrique. Un espace métrique est donc le couple constitué d'un ensemble E et d'une distance : (E, d). $\mathbb{R}^2$ donne donc naissance à trois espaces métriques : $(\mathbb{R}^2, d_1)$, $(\mathbb{R}^2, d_2)$, $(\mathbb{R}^2, d_3)$.

Distance de deux ensembles

Étant donné un ensemble E sur lequel est définie une distance d et deux sous-ensembles A et B de E, on appelle distance des deux ensembles A et B et on note d(A, B) la borne inférieure de l'ensemble des nombres $d(x, y)$ obtenus lorsque x varie dans A et y dans B.

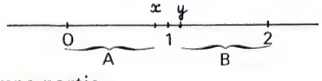
Par exemple, considérons les deux intervalles fermés A = [1, 2] et B = [3, 5] de $\mathbb{R}$. Si x est un élément quelconque de A et y un élément quelconque de B, la distance de x à y est $y - x$ (puisque y est supérieur à x). La plus petite des différences $y - x$ possibles est obtenue quand $x = 2$ et $y = 3$; la distance des deux ensembles A et B est, dans ce cas-là, 1.



Propriétés :

1) La distance de A à B est égale à la distance de B à A.

2) La distance de A à B peut être nulle sans que les ensembles A et B soient égaux : par exemple, si A est l'intervalle fermé [0, 1] et B l'intervalle fermé [1, 2] de $\mathbb{R}$, la distance $d(x, y)$ entre un point x de A et un point y de B est comprise entre 0 et 2 ($0 \leqslant d(x, y) \leqslant 2$). La borne inférieure de l'ensemble des nombres $d(x, y)$ est 0, donc d(A, B) = 0 et, pourtant, A n'est pas égal à B.



Distance d'un point à un ensemble

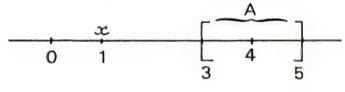
Soit E un ensemble muni d'une distance d. Si x est un point de E et A une partie de E, on appelle distance de x à A et on note $d(x, A)$ la borne inférieure des nombres $d(x, y)$ obtenus en faisant varier y dans A.

Exemple : dans $\mathbb{R}$, si $x = 1$ et A = [3, 5], $d(x, y) = y - 1$.
Le nombre $d(x, y)$ est compris entre $3 - 1$ et $5 - 1$:

$$2 \leqslant d(x, y) \leqslant 4.$$

La plus petite des distances possibles est donc 2 :

$$d(1, A) = 2.$$



Topologie d'un espace métrique ◆

◆ Lire d'abord Topologie.

Un espace métrique E peut toujours être muni d'une topologie : une partie A de E est un ouvert si et seulement si chaque point de A est centre d'une boule ouverte contenue dans A.
Pour cette topologie :
— une boule ouverte est un ouvert ;
— une boule fermée est un fermé ;
— une sphère est un fermé.
Les ouverts de cette topologie sont les réunions de boules ouvertes.

DISTRIBUTION (Loi de) : voir l'article « les Statistiques », page 491.

DISTRIBUTIVITÉ ◆ pages 26 et 27, 58, 73 à 75, 86, 267, 270, 305.

◆ Lire d'abord Composition (loi de) et l'article « l'Algèbre », pages 17 et suivantes.

Distributivité d'une loi de composition interne par rapport à une autre.

L'ensemble $\mathbb{N}$ des entiers naturels (positifs) est muni de deux lois de composition interne : l'addition notée + et la multiplication notée $\times$. Si on multiplie par le nombre a la somme $b + c$ (c'est-à-dire si on calcule l'élément $a \times (b + c)$), on trouve le même résultat que si on avait fait la somme des deux produits $a \times b$

et $a \times c$ (c'est-à-dire que si on avait calculé $(a \times b) + (a \times c)$), et cela est vrai, quels que soient les éléments a, b, c :

$$a \times (b + c) = (a \times b) + (a \times c).$$

La multiplication étant commutative, on a également :

$$(b + c) \times a = (b \times a) + (c \times a).$$

On dit que la multiplication est distributive par rapport à l'addition, dans l'ensemble des entiers naturels (positifs) $\mathbb{N}$.

De façon générale, si un ensemble E est muni de deux lois de composition interne notées $\divideontimes$ et T, on dit que la loi $\divideontimes$ est distributive par rapport à la loi T si elle est à la fois :

1) distributive à gauche par rapport à T ; c'est-à-dire si, quels que soient les éléments a, b, c de E, on a :

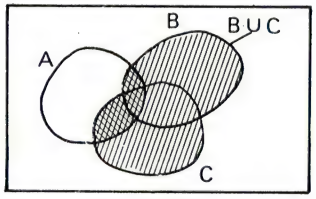
$$a \divideontimes (b \mathrel{T} c) = (a \divideontimes b) \mathrel{T} (a \divideontimes c).$$

2) distributive à droite par rapport à T ; c'est-à-dire si, quels que soient les éléments a, b, c de E, on a :

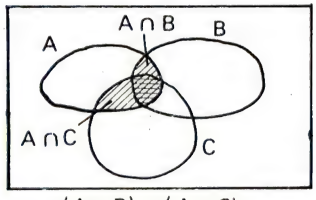
$$(b \mathrel{T} c) \divideontimes a = (b \divideontimes a) \mathrel{T} (c \divideontimes a).$$

Il est clair que, si la loi $\divideontimes$ est commutative, si elle est distributive par rapport à T à gauche ou à droite, elle est aussi distributive par rapport à T de l'autre côté et elle est donc distributive.
Nous avons vu que la multiplication est distributive par rapport à l'addition dans $\mathbb{N}$, mais l'addition n'est pas distributive par rapport à la multiplication ; en effet, $1 + (3 \times 2)$ n'est pas égal à $(1 + 3) \times (1 + 2)$ (car 7 n'est pas égal à 12).
Dans l'ensemble $\mathfrak{P}(E)$ des parties d'un ensemble E, la loi intersection $\cap$ est distributive par rapport à la loi réunion $\cup$, car, quels que soient les sous-ensembles A, B, C de E :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ (et la loi } \cap \text{ est commutative).}$$



A∩(B∪C)



(A∩B)∪(A∩C)

et la loi réunion est aussi distributive par rapport à la loi intersection :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \text{ (et la loi } \cup \text{ est commutative).}$$

DIVERGENCE : voir l'article « l'Analyse », page 109 et les pages 10, 11.

DIVISEUR DE ZÉRO page 48.

Si on considère les deux matrices carrées d'ordre 2 suivantes :

$$M = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \qquad N = \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$$

aucune de ces deux matrices n'est égale à la matrice nulle $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ et, pourtant, leur produit :

$$M \times N = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$$

est nul. On dit que les deux matrices M et N sont des diviseurs de zéro.
De façon générale, si A est un anneau unitaire, qui contient d'autres éléments que l'élément 0 (élément neutre de l'addition), un élément non nul a de A est un diviseur de 0 s'il existe un autre élément non nul b de A tel que le produit $a \times b$ soit nul ◆.

◆ Un anneau sans diviseurs de zéro est appelé un anneau intègre (ou anneau d'intégrité).

DIVISION pages 368 à 389.

Division exacte

Si un ensemble E est muni d'une loi de composition interne notée $\times$ (appelée multiplication), la division est l'opération inverse de la multiplication : étant donné deux éléments a et b de E, faire la division de a par b, c'est trouver un élément x de E tel que a soit égal au produit de b par x : $a = b \times x$.

Si x existe, on note $x = a : b$ ou $x = \frac{a}{b}$. Dans l'ensemble $\mathbb{N}^*$ des entiers naturels non nuls comme dans l'ensemble $\mathbb{Z}^*$ des entiers relatifs non nuls, cette division n'est pas toujours possible (ainsi, il n'est pas possible de trouver un entier naturel dont le produit par 2 soit égal à 5) ; par contre, elle l'est toujours dans $\mathbb{Q}^*$, ensemble des nombres rationnels non nuls.

Division à une unité près ou « division euclidienne »

Étant donné deux entiers naturels non nuls a et b, il existe deux autres entiers naturels q et r (le nombre r étant supérieur ou égal à 0 et inférieur à b) tels que le nombre a soit égal à la somme de r et du produit $b \times q$:

$$\begin{cases} a = b \times q + r \\ 0 \leqslant r < b. \end{cases}$$

q est appelé quotient et r reste de cette opération.

Exemple : $17 = 5 \times 3 + 2$; dans la division euclidienne de 17 par 5, le quotient est 3 et le reste 2.

Division euclidienne de deux polynômes (à coefficients réels)

Étant donné deux polynômes A et B, le polynôme B étant différent de 0, il existe deux autres polynômes Q et R ◆, R étant soit nul, soit de degré inférieur au degré de B, tels que A soit égal à la somme de R et du produit B × Q.

◆ Q et R déterminés de manière unique.

$$\begin{cases} A = B \times Q + R \\ R = 0 \text{ ou degré de } R < \text{degré de } B. \end{cases}$$

Trouver les polynômes Q et R, c'est effectuer la division euclidienne de A par B (ou encore division de A par B ordonnés selon leurs puissances décroissantes). Q s'appelle le quotient et R le reste. Une telle opération peut s'effectuer exactement comme une division ordinaire.

Exemple

$A = x^3 + x^2 - 2x + 1$
$B = x + 3$

Disposition pratique du calcul :

$$\begin{array}{r|l} x^3 + x^2 - 2x + 1 & x + 3 \\ \cline{2-2} -(x^3 + 3x^2) & \underbrace{x^2 - 2x + 4}_{Q} \\ \cline{1-1} -2x^2 - 2x + 1 & \\ -(-2x^2 - 6x) & \\ \cline{1-1} +4x + 1 & \\ -(4x + 12) & \\ \cline{1-1} \underbrace{-11}_{R} & \end{array}$$

$$\underbrace{x^3 + x^2 - 2x + 1}_{A} = \underbrace{(x + 3)}_{B} \times \underbrace{(x^2 - 2x + 4)}_{Q} \underbrace{- 11}_{R}$$

Division de deux polynômes suivant les puissances croissantes
Étant donné un polynôme A, un polynôme B dont le terme constant ne soit pas nul et un nombre entier naturel *n* strictement positif, il existe un polynôme Q et un polynôme R (déterminés de manière unique) tels que :

$\begin{cases} A = B \times Q + x^{n+1}.R \\ \text{avec degré de Q inférieur ou égal à } n. \end{cases}$

Trouver Q et R, c'est effectuer la division de A par B, suivant les puissances croissantes et à l'ordre *n*.

Exemple

$A = 1 + x$
$B = 1 - x$

effectuer la division suivant les puissances croissantes de A par B à l'ordre 2.

$$\begin{array}{r|l} 1 + x & 1 - x \\ \cline{2-2} -(1 - x) & \underbrace{1 + 2x + 2x^2}_{Q} \\ \cline{1-1} 2x & \\ -(2x - 2x^2) & \\ \cline{1-1} 2x^2 & \\ -(2x^2 - 2x^3) & \\ \cline{1-1} 2x^3 & \end{array}$$

$$1 + x = (1 - x)\underbrace{(1 + 2x + 2x^2)}_{Q} + \underbrace{x^3 \times 2}_{R}$$

Division harmonique : voir Birapport.

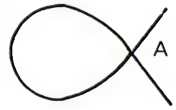
DODÉCAÈDRE : voir Polyèdres

DODÉCAGONE : voir Polygones

DOUBLE pages 260, 444.

Point double d'une courbe

Soit (C) une courbe plane définie par ses équations paramétriques $x = f(t)$ et $y = g(t)$. La courbe présente un point double s'il existe *deux valeurs distinctes* t_1 et t_2 du paramètre t pour lesquelles $f(t_1) = f(t_2)$ et $g(t_1) = g(t_2)$.



Par exemple, la courbe d'équations : $\begin{cases} x = t^2 + 1 \\ y = t^3 - t + 1 \end{cases}$ admet un point double.

Car, si $t = 1$, alors $x = 2$ et $y = 1$

si $t = -1$, on a aussi $x = 2$ $y = 1$.

En général, la courbe admet deux tangentes distinctes au point A (une pour chacun des arcs qui se coupent en A).

Point double d'une transformation géométrique

Tout point qui, dans cette transformation, a lui-même pour image. Par exemple, dans une symétrie par rapport à un point O, O est le seul point double ; dans une symétrie par rapport à une droite D, tout point de la droite D est point double.

DROITE AFFINE, VECTORIELLE : voir l'article « l'Algèbre linéaire », pages 63, 70, 71.

DROITE DE RÉGRESSION : voir l'article « les Statistiques », pages 476 à 480.

DROITE NUMÉRIQUE ◆ pages 13, 216.

◆ Voir le mot Topologie.

Ensemble ordonné des nombres réels $\mathbb{R}$, muni de la topologie de l'ordre.

Droite numérique achevée

Si on adjoint à $\mathbb{R}$, muni de la relation d'ordre $\leqslant$, les deux éléments notés $-\infty$ et $+\infty$ ◆ satisfaisant aux propriétés suivantes :

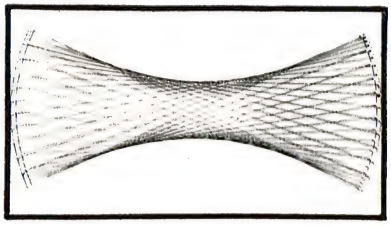
◆ Moins l'infini et plus l'infini.

1) $-\infty < +\infty$

2) quel que soit x, appartenant à $\mathbb{R}$, on a :

$$-\infty < x \text{ et } x < +\infty$$

on obtient un ensemble, totalement ordonné, noté $\overline{\mathbb{R}}$ (« $\mathbb{R}$ barre ») appelé droite numérique achevée.



E

e : voir Logarithme népérien

ÉCART, ÉCART TYPE, ÉCART RÉDUIT :
voir l'article « les Statistiques », pages 467 à 471, 496, 498, 508, 510, et les pages 429 à 431.

ÉCHANTILLONNAGE : voir l'article « les Statistiques », pages 457, 509 à 518.

ÉGALITÉ page 395.

Deux objets mathématiques a et b sont égaux si a et b sont deux représentations différentes du *même* objet. L'égalité mathématique est, en fait, l'identité. Les figures égales dont on parlait autrefois n'ont rien à voir avec cette notion d'égalité. Ces figures sont des figures isométriques (ou équivalentes).

ÉLÉMENT NEUTRE◆ pages 22 et 23, 56, 85.

◆ Lire aussi l'article « l'Algèbre », page 22.

Expression à prendre au sens propre. Prenons un exemple dans l'ensemble des nombres entiers positifs $\mathbb{N}$:

$n + 0 = 0 + n = n$ 0 est l'élément neutre de l'addition

$n \times 1 = 1 \times n = n$ 1 est l'élément neutre de la multiplication.

Soit E un ensemble muni de la loi de composition interne notée $\divideontimes$; s'il existe un élément e' de E tel que, pour tout x de E, le composé de e' avec x soit égal à x ($e' \divideontimes x = x$), l'élément e' est dit élément neutre à gauche pour la loi $\divideontimes$; s'il existe un élément e'' de E tel que, pour tout x de E, le composé de x avec e'' soit égal à x ($x \divideontimes e'' = x$), l'élément e'' est dit élément neutre à droite pour la loi $\divideontimes$.
S'il existe à la fois un élément neutre à gauche et un élément neutre à droite dans l'ensemble E pour la loi $\divideontimes$, ces deux éléments sont égaux.
On dit qu'un élément e d'un ensemble E, muni d'une loi de composition interne $\divideontimes$, est élément neutre pour la loi $\divideontimes$, s'il est à la fois élément neutre à gauche et élément neutre à droite.

Exemple

Dans $\mathbb{N}$ (entiers naturels positifs), la loi exponentiation admet un élément neutre à droite : 1, puisque, pour tout entier n de $\mathbb{N}$:

$n \divideontimes 1 = n^1 = n.$

Mais cet élément n'est pas neutre à gauche, car, pour tout n de $\mathbb{N}$:

$$1 \divideontimes n = 1^n = 1 \text{ (et non pas } n).$$

De même dans $\mathbb{Z}$ (ensemble des entiers relatifs, positifs et négatifs), la soustraction admet un élément neutre à droite : 0, car, pour tout n de $\mathbb{Z}$, $n - 0 = n$, mais cet élément neutre à droite n'est pas élément neutre à gauche, car $0 - n = -n$ et non pas n.

Remarque : si la loi $\divideontimes$ est commutative, un élément neutre à droite est aussi élément neutre à gauche, donc élément neutre. Dans $\mathbb{N}$, 0 est élément neutre pour l'addition, 1 est élément neutre pour la multiplication.

ÉLÉMENT RÉGULIER◆ pages 25 et 26.

◆ Lire aussi l'article « l'Algèbre », page 25.

Dans l'ensemble des entiers naturels positifs $\mathbb{N}$:
— pour l'addition,

$$a + x = a + y \Longrightarrow x = y$$
$$x + a = y + a \Longrightarrow x = y$$

— pour la multiplication,

$$a \times x = a \times y \Longrightarrow x = y$$
$$x \times a = y \times a \Longrightarrow x = y$$

à condition que a soit différent de 0 (si $a = 0$, l'égalité $0 \times x = 0 \times y$ a lieu même si x est différent de y). Un élément a d'un ensemble E muni d'une loi de composition interne notée $\divideontimes$ est dit régulier à gauche si l'égalité $a \divideontimes x = a \divideontimes y$ implique $x = y$. Il est dit régulier à droite si l'égalité $x \divideontimes a = y \divideontimes a$ implique $x = y$; un élément a est dit régulier s'il est à la fois régulier à gauche et à droite.

Exemple : dans $\mathbb{N}$, nous l'avons vu, tout élément est régulier pour l'addition, mais, pour la multiplication, 0 n'est régulier ni à gauche ni à droite.

ÉLÉMENT SYMÉTRIQUE◆ pages 23 et 24.

◆ Lire aussi l'article « l'Algèbre », page 23.

Expression à prendre au sens propre. Prenons un exemple dans l'ensemble des rationnels $\mathbb{Q}$:

$$2 + (-2) = 0 \qquad 5 \times \frac{1}{5} = 1$$

— 2 est le symétrique de 2 pour l'addition ; $\frac{1}{5}$ est le symétrique de 5 pour la multiplication, car la somme de 2 et de — 2 est égale à l'élément neutre◆ 0 de l'addition et le produit de 5 par $\frac{1}{5}$ est égal à l'élément neutre 1 de la multiplication.

◆ Voir ce mot.

Si E est un ensemble muni d'une loi de composition interne notée $\divideontimes$ pour laquelle il existe un élément neutre e, l'élément x' de E est dit symétrique à gauche de l'élément x de E si le composé de x' et x est e : $x' \divideontimes x = e$.
Un élément x'' de E est dit symétrique à droite de x si le composé de x et x' est e : $x \divideontimes x'' = e$ (si la loi $\divideontimes$ est associative et si x admet un symétrique à droite : x'', et un symétrique à gauche : x', ces deux éléments sont égaux). Un

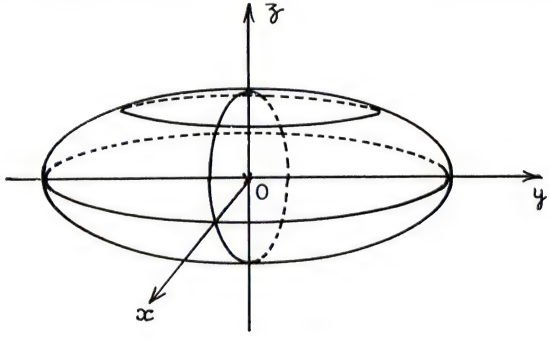
élément x' de E est dit symétrique de l'élément x de E si on a à la fois $x' \ast x = e$ et $x \ast x' = e$; si la loi $\ast$ est associative, le symétrique x' de x est unique (notons que, si x' est symétrique de x, x est symétrique de x' ; on dit simplement que x et x' sont symétriques l'un de l'autre). Si la loi de composition interne est notée $+$ le symétrique de x, s'il existe, est noté $-x$ et il est appelé l'opposé de x ; si la loi est notée $\times$, le symétrique de x, s'il existe, est noté x^{-1} ou $\frac{1}{x}$ et il est appelé l'inverse de x.

ELLIPSE : voir Coniques et la page 397.

ELLIPSOÏDE

Surface dont l'équation est du second degré et dont toutes les sections planes sont des ellipses ou des cercles. Il y a trois axes de symétrie perpendiculaires deux à deux ; leur intersection est appelée centre de l'ellipsoïde. Si ces trois axes de symétrie sont les trois axes d'un système de coordonnées, l'équation de l'ellipsoïde est de la forme :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0.$$



(Dans le cas où les trois nombres a, b, c sont égaux, l'ellipsoïde est une sphère.)
Un ellipsoïde de révolution est un ellipsoïde dont les sections par les plans perpendiculaires à l'un des axes sont des cercles. Un tel ellipsoïde peut être engendré par rotation d'une ellipse autour de l'un de ses axes.

ENDOMORPHISME : voir Homomorphisme et la page 229.

ENNÉAGONE : voir Polygones

ENSEMBLE : voir l'article page suivante et les pages 167, 173, 174, 191, 203, 246, 308, 330, 335, 336, 361, 375 à 389, 394, 395, 398, 445, 446, 527, 533, 534.

Les ensembles

par Pierre Buisson

Dans le langage courant, ce mot est employé avec diverses significations comme synonyme de groupement, collection, classe, etc. Il a par contre, en mathématique, un sens très précis ; le but de cette rubrique est de donner, par divers modes d'approche, les règles d'emploi permettant d'affirmer que ce que nous désignons par groupement, collection, classe, etc., a le droit de s'appeler ensemble.

De tout temps, les mathématiciens se sont intéressés à des objets mathématiques non pas étudiés pour eux-mêmes, mais regroupés dans ce que nous appellerons un ensemble, attachant moins d'importance aux objets qu'aux relations possibles entre ces objets regroupés en ensembles.

Toutes les théories mathématiques (algèbre, géométrie, topologie, probabilités...) ont en commun l'utilisation d'un même type de raisonnement et des résultats ne faisant intervenir que les ensembles et non les structures mises sur ces ensembles. Plus précisément, nous pouvons dire que toutes ces théories sont contenues dans la logique et, sauf éventuellement l'arithmétique, dans la théorie des ensembles (par « contenus », nous entendons : tout théorème de la théorie des ensembles est un théorème de toute autre théorie). L'utilisation explicite de ces théorèmes commence seulement à faire son apparition dans l'enseignement ; remarquons d'ailleurs qu'ils n'ont été clairement formalisés par les mathématiciens que depuis le début de ce siècle.

Les mathématiciens construisirent les nombres négatifs, rationnels, réels et complexes à partir des entiers naturels ; mais ils laissèrent à Dieu la construction de ces derniers en disant qu'ils étaient dans la nature (Kronecker : les nombres entiers ont été faits par le Bon Dieu, tous les autres sont de fabrication humaine). Pourtant, là aussi, les ensembles permettent de dégager la notion de nombre entier ou

Pierre Buisson
Né en 1940 ; diplômes d'enseignement supérieur, puis doctorat de troisième cycle à Strasbourg. Maître-assistant à l'université de Rouen ; poursuit à l'Institut de recherches sur l'enseignement des mathématiques de Rouen une étude sur cet enseignement dans le premier cycle de l'enseignement secondaire : analyse des difficultés, investigation des possibilités intellectuelles des élèves, etc.

entier naturel, et cela, semble-t-il, de manière pédagogiquement efficace pour les jeunes enfants.

Nous pouvons dégager quatre modes principaux d'appréhension des ensembles. Tout d'abord, un mode expérimental consistant, par des manipulations ensemblistes, à constater certains résultats, à les traduire clairement, à les utiliser ensuite pour la compréhension d'autres notions. Cette manipulation peut se faire à deux niveaux : niveau concret, avec des objets que l'on « prend en main » (à l'école maternelle, au cours préparatoire) ; niveau abstrait, avec des objets conçus par la pensée (début du premier cycle de l'enseignement secondaire). Puis, supposant faites ces manipulations, un mode analytique permettant de dégager, à partir des résultats expérimentaux, d'autres résultats plus complexes et de les utiliser dans les différentes branches de la mathématique. Jusque-là, nous parlerons de langage ensembliste et non pas de théorie des ensembles. Quoi que l'on puisse lire, il n'a jamais été question d'introduire la théorie des ensembles dans une classe de l'enseignement primaire ou secondaire. Un troisième mode consiste à faire une théorie « naïve » des ensembles ◆. Théorie, car les axiomes sont clairement précisés et les démonstrations effectives ; naïve, car en partie seulement formalisée. Enfin, un dernier mode consistant en une théorie axiomatique formalisée ◆, dans laquelle il n'est plus fait appel à l'intuition, et réservée aux spécialistes.

◆ Une « théorie naïve » est d'abord la donnée d'un certain nombre de phrases écrites en langage ordinaire, ayant un sens et prises pour vraies en référence à une certaine intuition ; ce sont les axiomes de la théorie. On développe alors cette théorie en posant des définitions et en déduisant d'autres résultats à partir de ces définitions et des axiomes. Le lecteur qui voudra connaître les axiomes d'une théorie des ensembles et les principaux résultats de celle-ci pourra se reporter au livre de R. Halmos : *Introduction à la théorie des ensembles* (Paris, Gauthier-Villars). La connaissance d'une telle théorie naïve n'est pas nécessaire à celle des autres rubriques de ce dictionnaire.

◆ Une théorie axiomatique formalisée utilise un langage formel et les mots du langage courant comme « appartenir » ou « entier », qui désignent certaines formules, peuvent ne plus avoir leur signification intuitive. Cette théorie est exposée en détail dans *la Théorie axiomatique des ensembles* de J.-C. Krivine (Paris, P.U.F., coll. « SUP »).

Dans ce qui va suivre, il ne sera pas question de manipulations concrètes. Une première partie montrera comment, de la notion d'ensemble et des opérations ensemblistes, on peut dégager la notion de nombre entier, ainsi que les opérations sur ces nombres. Nous y ferons largement appel à l'intuition du lecteur. Une seconde partie, indispensable à la compréhension des autres rubriques, sera consacrée à l'étude des relations.

ENSEMBLES ET NOMBRES ENTIERS

Un discours mathématique cohérent nécessite des phrases ayant un sens non seulement en français, mais encore en mathématique. Pour éviter le cercle vicieux qui consiste à définir certains mots à l'aide d'autres mots, et ces derniers à l'aide des premiers, nous sommes amenés à utiliser des mots de base, appelés « mots premiers », que nous ne définissons pas ; et c'est à partir de ces mots premiers que nous pourrons définir d'autres mots.

Ensemble, élément, appartient

Voilà trois mots premiers qui, pris isolément, n'ont d'ailleurs aucun sens mathématique ; seule la phrase « l'élément *a* appartient à l'ensemble A » a un sens ; elle se note :

$$a \in A$$

Une collection de timbres, une classe dans une école, une famille, un alphabet peuvent donner une idée concrète de ce qu'est, en mathématique, un ensemble d'éléments ; cela toutes les fois où les objets, ainsi réunis en un tout, peuvent être perçus séparément par nos sens ou notre pensée, cette perception étant indépendante de celui qui perçoit, donc identique pour tous. Les personnes blondes ne peuvent être regroupées dans un ensemble, car il n'existe pas de critère universellement reconnu caractérisant la blondeur. Afin de pouvoir concrétiser la notion d'ensembles, il faut une grande précision de langage. Ainsi, nous pouvons parler de l'ensemble des élèves d'une classe présents à un instant précis ou de l'ensemble des élèves inscrits dans cette classe ; mais il est dangereux de parler de l'ensemble des élèves de la classe sans donner d'autres précisions. Un tas de cailloux, un sac de pommes de terre ne concrétisent pas la notion d'ensemble, car les objets qui sont regroupés ne se distinguent pas clairement les uns des autres.

Nous admettrons donc que nous avons affaire à un ensemble d'éléments si chaque élément est clairement défini, distinguable de tout autre, et si l'appartenance de l'élément à l'ensemble est claire pour tout le monde.

L'exemple le plus simple est la donnée d'un ensemble par la liste exhaustive de ses éléments. On dit que l'ensemble est défini en extension, et on le note avec des accolades ; par exemple, l'ensemble dont les éléments sont les dessins O, ✲ et × se note $\{ \text{O}, ✲, \times \}$. Des figures géométriques donnent un bon exemple d'ensemble, à condition que chaque figure se distingue de n'importe quelle autre par au moins une propriété (couleur, forme, épaisseur...). Il suffit donc, pour construire un ensemble en extension, de disposer d'objets clairement identifiables ; nous avons ainsi à notre disposition de nombreux ensembles que nous pouvons décrire et manier.

Exemple. Si la lettre *a* désigne un objet, il est naturellement possible de définir l'ensemble dont l'objet désigné par *a* est l'unique élément. Cet ensemble noté $\{ a \}$ s'appelle un singleton, et l'on a : $a \in \{ a \}$.

Égalité

En mathématique, l'égalité est l'identité ; cela est simple à dire mais difficile à expliquer. Les objets qui sont perçus par nos sens ou notre pensée sont des signifiés ; les signes et les mots utilisés pour les écrire sont des signifiants et désignent les objets. Par exemple, le chiffre 2

désigne l'objet conçu par notre pensée qui est le nombre deux. Un même objet peut avoir plusieurs signifiants distincts ; le nombre deux peut aussi être désigné par « 1 + 1 » ; on écrira alors « 1 + 1 = 2 », qui se lit : « Un plus un égale deux ». Un objet que l'on « prend en main » ne peut être égal qu'à lui-même ; il est donc impossible de concevoir concrètement l'égalité de deux objets. En général, les objets sont désignés par des lettres, des chiffres ou des signes, mais il peut arriver que l'on veuille considérer ces symboles en tant qu'objet ; il vaut alors mieux les mettre entre guillemets. Avec cette convention, un ensemble de lettres devra s'écrire par exemple $\{$ « a », « z », « e » $\}$.

Nous avons admis implicitement qu'un ensemble était déterminé de manière unique par la donnée de ses éléments ; il ne peut y avoir d'autre ensemble ayant les mêmes éléments. L'ensemble désigné par E est égal à l'ensemble désigné par F si tout élément de E est élément de F et réciproquement ; nous écrivons alors E = F. Nous utiliserons le signe « = » (qui ne s'introduit qu'au niveau des signifiants) sans aucune considération « philosophique » ; nous reviendrons sur le concept d'égalité lors de l'étude des *ensembles quotients*.
Il est indispensable de bien faire cette distinction entre signifiant et signifié afin de traduire correctement certaines phrases incorrectes : « Soient A un ensemble et a un élément de A » donnera : « Soient un ensemble désigné par A et un élément de cet ensemble désigné par a. » Souvent, nous écrirons un énoncé du premier type, laissant à chacun le soin de reconnaître un tel abus de langage.

Propriété sur un ensemble

Soit E l'ensemble, donné en extension, des membres d'une famille. Si x est un élément de E, la phrase « x est un homme » est soit vraie, soit fausse. Nous pouvons ainsi définir un ensemble H : l'ensemble dont les éléments sont ceux de l'ensemble E pour lesquels la phrase est vraie ; on notera ◆ :

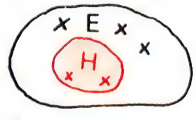
$$H = \{\, x \in E \mid x \text{ soit un homme} \,\}$$

(on lit : ensemble des x appartenant à E tels que x soit un homme). Par construction, tous les éléments de H sont éléments de E ; on dit alors que H est un *sous-ensemble* ou une *partie* de E, ce qui s'écrit :

$$H \subset E$$ (et se lit : H inclus dans E).

Si H et E sont distincts, on dit que H est un *sous-ensemble strict* de E. D'une manière générale, nous appellerons propriété P(x) ◆ sur un ensemble E la donnée d'une phrase (« x est un homme », dans l'exemple) où figure la lettre x appelée « variable » ; cette phrase,

◆ Le signe | signifie « tel que ».



H est un sous-ensemble *strict* de l'ensemble E.

◆ La lettre x peut être remplacée par n'importe quelle autre lettre comme y, a, b, ..., ou par un espace blanc. C'est pourquoi il faudrait dire « une propriété P » et non « une propriété P(x) ; la lettre x ne joue ici qu'un rôle accessoire.

notée P(x), doit avoir un sens chaque fois que x est remplacé par n'importe quel élément de E, et elle doit être alors soit vraie, soit fausse. Nous pouvons alors considérer l'ensemble F des éléments de E qui vérifient la propriété P(x) ; nous écrivons $F = \{ x \in E \mid P(x) \}$ que nous lisons : *ensemble des éléments x appartenant à* E *et vérifiant* P(x) [on dit aussi par abus de langage : ensemble des x de E tels que P(x)◆]. L'ensemble F est donc clairement défini par la donnée de l'ensemble E et de la propriété P(x) ; on dit alors que l'ensemble F est défini en compréhension. En fait, nous ne définissons ainsi que des sous-ensembles.

◆ On peut aussi lire, sans utiliser la lettre x : « ensemble des éléments de E vérifiant P ».

Propriétés équivalentes

Nous dirons que deux propriétés définies sur le même ensemble sont équivalentes si et seulement si elles définissent le même sous-ensemble. Par exemple P(x) et Q(x) sont équivalentes si et seulement si les sous-ensembles $F = \{ x \in E \mid P(x) \}$ et $G = \{ x \in E \mid Q(x) \}$ sont les mêmes, c'est-à-dire $F = G$.

Il est souvent utile de concrétiser des ensembles par des schémas qui jouent un rôle analogue aux « figures » de la géométrie traditionnelle. Si a, b, c désignent des objets distincts, on illustre l'ensemble $E = \{ a, b, c \}$ en traçant une courbe fermée sans points doubles dans le plan et en figurant les éléments à l'intérieur de cette courbe. Un tel dessin s'appelle diagramme de Venn ou d'Euler. Ces diagrammes peuvent être utilisés dans des exercices d'initiation, mais il faut prendre garde à ne pas en abuser, car, mal utilisés, ils peuvent être dangereux.

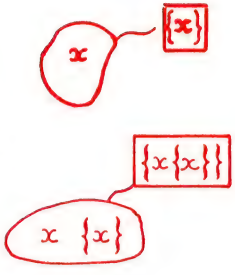
À un niveau plus avancé, ils sont utiles pour dégager certaines propriétés générales des opérations ensemblistes. Nous pourrons ainsi représenter un ensemble quelconque par une partie du plan intérieure à une courbe fermée sans point double (un cercle, un rectangle). On utilise alors ces diagrammes comme support de l'intuition.

Les limites d'un exposé intuitif

L'on convient généralement de désigner par des lettres majuscules les ensembles et par des lettres minuscules les éléments ; mais cela n'est pas toujours possible et, de plus, il ne faut pas en conclure que certains objets ne peuvent avoir que le statut d'éléments et d'autres celui d'ensembles : une droite peut être considérée à la fois comme ensemble de points et comme éléments de l'ensemble des droites du plan.

Aussi faut-il se placer dans un univers d'objets dont certains peuvent, selon les circonstances, être considérés comme élément ou comme ensemble. Ces objets ne nous sont pas tous donnés à priori, mais nous pouvons en construire indéfiniment.

Exemple



À partir d'un objet noté x, nous pouvons construire le singleton $\{ x \}$, qui est un objet, un ensemble, distinct de x. Nous pouvons alors construire d'autres ensembles en extension :

$\{ \{ x \} \}$, $\{ x, \{ x \} \}$, etc. (on lit pour $\{ \{ x \} \}$: ensemble formé du singleton $\{ x \}$; $\{ x, \{ x \} \}$ est l'ensemble formé de x et du singleton $\{ x \}$).

Conservant le point de vue intuitif, nous serions tentés de dire que les ensembles sont des objets clairement identifiables et d'en conclure qu'il existe un ensemble de tous les ensembles. Or supposons qu'un tel ensemble existe ; notons-le E ; sur E, considérons la propriété « x n'appartient pas à x », que nous notons ainsi : « $x \notin x$ ».

Prenons un exemple. Dans une bibliothèque, le bibliothécaire décide de faire établir pour chaque rayon un catalogue, destiné à figurer dans le rayon. Certains employés, comme il n'y a pas de consigne à ce sujet, mentionnent le catalogue dans le catalogue, d'autres non. Le bibliothécaire veut alors établir le catalogue de tous les catalogues qui ne se mentionnent pas, « x n'appartient pas à x ».
Puis il se demande si ce dernier catalogue doit se mentionner ou non. Si ce catalogue se mentionne, cela voudrait dire qu'il ne se mentionne pas (puisqu'il serait dans la liste des catalogues ne se mentionnant pas) ; ce qui est une absurdité. S'il ne se mentionne pas, cela voudrait dire qu'il devrait appartenir à la liste des catalogues ne se mentionnant pas, donc se mentionner.
Conclusion : on ne peut pas faire un tel catalogue.

Selon le même raisonnement, on ne peut obtenir un ensemble de tous les ensembles.

Il était possible de dresser la liste de tous les catalogues ne se mentionnant pas, mais nous ne pouvons pas lui donner le nom de « catalogue », sous peine d'absurdité ; nous pouvons l'appeler « liste », qui n'a pas la propriété de « catalogue ». De même, nous parlerons de la collection des ensembles et de la collection des ensembles qui ne s'appartiennent pas ; parler de l'ensemble des ensembles qui ne s'appartiennent pas aboutirait à une contradiction. En effet, soit A le sous-ensemble associé à la propriété « $x \notin x$ » ; $A = \{ x \in E \mid x \notin x \}$; si $A \in A$, alors par définition de A on a $A \notin A$ et, si $A \notin A$, alors $A \in A$; nous devons donc avoir en même temps $A \in A$ et $A \notin A$, ce qui est contradictoire. Ce paradoxe est dû à Russell ◆.

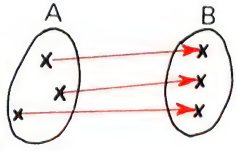
◆ B. Russell (1872-1970) : philosophe, logicien et mathématicien anglais dont l'œuvre immense porte aussi bien sur l'éthique et la politique que sur les fondements de la mathématique et de la pensée scientifique.

Nous dirons que toute propriété définit une collection ; lorsque cette collection est un ensemble, nous dirons que la propriété est collectivisante.

DES ENSEMBLES AUX NOMBRES

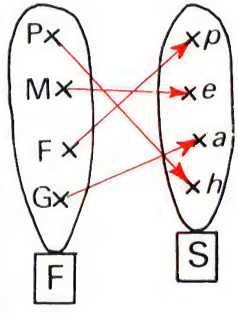
Nombres sans compter

Considérons une famille à table, cette famille étant composée du père, de la mère, d'un garçon et d'une fille ; nous avons ainsi un ensemble F. Désignons par E l'ensemble des enveloppes de serviette, un signe distinctif, un dessin par exemple, permettant d'attribuer à chaque membre de la famille une enveloppe et une seule, toutes étant ainsi réparties et chacune à une personne différente. Nous pouvons alors affirmer, sans avoir à compter, qu'il y a autant d'enveloppes de serviette que de personnes à table ; nous dirons que nous établissons, par cette attribution, une correspondance terme à terme ou bijection de l'ensemble F sur l'ensemble E. Lorsque, dans une réunion, toutes les personnes sont assises et qu'aucun siège ne reste inoccupé, nous pouvons affirmer sans avoir à compter qu'il y a autant de sièges que de personnes, car il y a une manière naturelle d'établir une bijection de l'ensemble des sièges sur l'ensemble des personnes présentes.

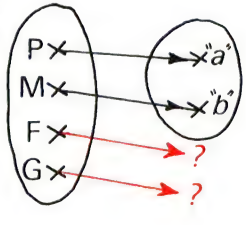


La bijection. D'une manière générale, si, étant donné deux ensembles A et B, nous pouvons faire correspondre à tout élément de A un élément de B et un seul, chaque élément de B étant ainsi associé à un élément de A et à un seul, nous dirons que nous avons établi une bijection de l'ensemble A sur l'ensemble B ; une expression synonyme est alors : les ensembles A et B ont même nombre d'éléments. Nous dirons alors que A et B sont équipotents.

La notion de classe. Les bijections nous permettent de faire une classification des ensembles, les ensembles A et B étant dans la même classe si et seulement s'il est possible d'établir une bijection de A sur B (il est d'ailleurs alors possible d'établir une bijection de B sur A). Un ensemble ne pourra se trouver que dans une seule classe Dans des cas simples, on peut opérer à l'aide des diagrammes de la façon suivante : de chaque élément d'un ensemble on fait partir une flèche et une seule aboutissant à un élément de l'autre ensemble ; il y aura bijection si chaque élément du dernier ensemble est atteint une et une seule fois.



Si $F = \{ P, M, G, F \}$ est l'ensemble familial précédemment décrit (père, mère, garçon, fille) et $S = \{ p, e, a, h \}$ l'ensemble des saisons (printemps, été, automne, hiver), nous remarquons qu'il existe une bijection entre F et S (de chaque élément de F part une flèche et une seule aboutissant à un élément et un seul de S et tout élément de S est atteint une et une seule fois).



Les ensembles F et S appartiennent à la même classe suivant la définition précédente. Par contre, il n'existe pas de bijection entre l'en-

semble F et l'ensemble $\{$ « a », « b » $\}$. Ils ne sont pas dans la même classe.

Notion de nombre

Les notions d'ensembles et de bijections permettent de dégager la notion de nombres indépendamment de toute dénomination et de toute écriture. Poser la question : « Quel est le nombre d'éléments de l'ensemble A ? » revient à demander : « Dans quelle classe d'ensembles se trouve A ? »

Nombres en comptant. Pour répondre à la question précédente, il nous faut d'abord donner un nom à chaque nombre, c'est-à-dire à chaque classe d'ensembles. Il est clair que tous les singletons sont dans la même classe, nous la nommons « un » ; la classe où figurent les ensembles F et S des exemples précédents s'appelle « quatre »... Il nous faut ensuite choisir dans chaque classe un ensemble de référence.
Vient donc en premier lieu la notion de nombres auxquels nous avons ensuite donné des noms ; enfin seulement nous allons les écrire en les symbolisant chacun par un caractère. Habituellement, le symbole du nombre « un » est « 1 », celui de « deux » est « 2 », celui de « quatre » est « 4 » ; les Romains avaient d'ailleurs choisi d'autres symboles. Nous appelons chiffres les symboles 1, 2, 3, 4, etc. ; le problème de la numération est celui de l'écriture de chaque nombre à l'aide de chiffres ; c'est un problème délicat qu'il n'est pas possible d'aborder ici.

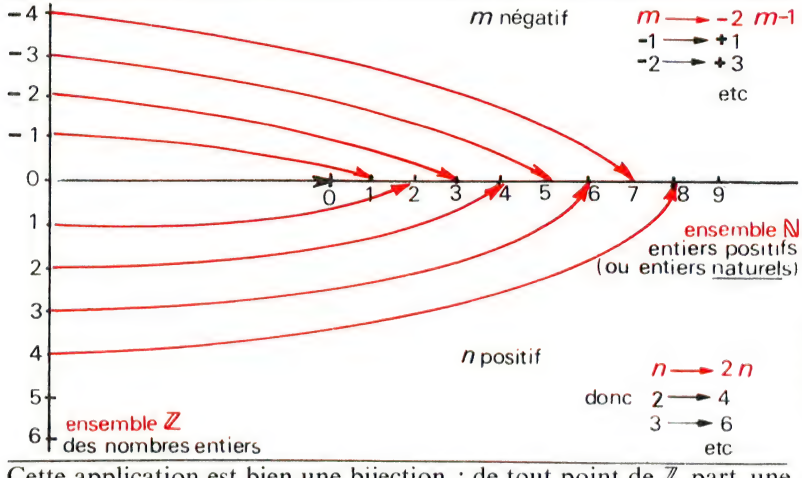
Ensemble vide et nombre zéro. Sur un ensemble A, prenons une propriété qui n'est vraie pour aucun de ses éléments, par exemple P(x) : « $x \neq x$ », x non égal à x. Nous définissons ainsi un sous-ensemble de A n'ayant aucun élément, appelé sous-ensemble vide de A ; nous notons $\varnothing_A = \{ x \in A \mid x \neq x \}$. Si B désigne un autre ensemble, nous pouvons de même définir le sous-ensemble vide de B, noté $\varnothing_B$. La logique permettra de comprendre le délicat raisonnement suivant : l'ensemble $\varnothing_A$ n'ayant aucun élément, la phrase « tout élément de $\varnothing_A$ est élément d'un ensemble E » est vraie, car il n'y a rien à vérifier ; nous avons donc pour tout ensemble E l'inclusion $\varnothing_A \subset E$, et en particulier $\varnothing_A \subset \varnothing_B$; de la même manière, nous démontrerions l'inclusion $\varnothing_B \subset \varnothing_A$; nous avons donc l'égalité $\varnothing_A = \varnothing_B$. Ces deux symboles désignent le même ensemble appelé ensemble vide et noté $\varnothing$. On appelle zéro, noté 0, le nombre d'éléments de l'ensemble vide.

Nombres sans pouvoir compter. Nous avons vu que nous pouvons définir une classification des ensembles, les ensembles A et B étant

dans la même classe, si et seulement s'il est possible d'établir une bijection de A sur B. Nous appelons cardinal de A, noté Card A, la classe dont fait partie l'ensemble A. Les deux symboles « Card A » et « Card B » désignent le même objet si et seulement si les ensembles A et B sont équipotents ◆ ; se donner un cardinal x revient à se donner un ensemble A tel que x = Card A. Nous attachons ainsi à chaque ensemble son cardinal, appelé aussi nombre d'éléments.

◆ Voir ci-dessus la définition de la bijection.

Cardinal fini et cardinal infini. Prenons un exemple. Nous allons construire une bijection de l'ensemble de tous les entiers (positifs et négatifs $\mathbb{Z}$) dans l'ensemble des entiers positifs seulement $\mathbb{N}$, ce qui démontrera que l'on peut réunir autant d'entiers positifs que d'entiers positifs et négatifs.



À 0 j'associe 0,
à tout nombre positif n j'associe le nombre positif $2n$,
à tout nombre négatif m j'associe le nombre positif $-2m-1$.

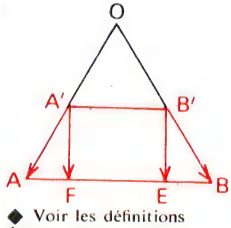
Cette application est bien une bijection : de tout point de $\mathbb{Z}$ part une flèche et une seule aboutissant à un élément de $\mathbb{N}$ et chaque élément de $\mathbb{N}$ est atteint une et une seule fois.

$\mathbb{N}$ est un sous-ensemble strict ◆ de $\mathbb{Z}$.

◆ Voir la définition d'un sous-ensemble strict, page 259.

Par contre, prenons un ensemble E de 5 points du plan et un sous-ensemble strict de E, noté F. F aura 1, 2, 3 ou 4 points mais pas 5, car il s'agit d'un sous-ensemble strict. Il n'est pas possible de définir une bijection de F dans E. Cette propriété est le contraire de la propriété précédente.

Prenons un exemple géométrique ; soient (O, A, B) un triangle isocèle, A′ et B′ respectivement les milieux des côtés [OA] et [OB]. Construisons alors le segment [EF] inclus dans le segment [AB] tel que le quadrilatère (A′, B′, E, F) soit un rectangle. La figure ci-contre indique clairement l'existence d'une bijection entre les segments [A′B′] et [FE], d'une part (translation), et les segments [A′B′] et [AB], d'autre part (homothétie) ◆ ; les trois segments



◆ Voir les définitions de ces mots.

[AB], [A′B′] et [EF] ont donc même nombre d'éléments (appelés ici points). En particulier, le segment [AB] a même nombre de points que le segment [EF], qui est un sous-ensemble strict.
Il est évidemment impossible de mettre de tels ensembles en bijection avec un ensemble de doigts, par exemple, car passer d'un ensemble de doigts à un autre en rajoutant un doigt change le nombre d'éléments, alors qu'agrandir un segment ne le change pas.
Nous appelons cardinal fini ou nombre entier ou entier naturel le nombre d'éléments d'un ensemble A *qui possède la propriété suivante :* « *Si* A′ *est un sous-ensemble strict (c'est-à-dire distinct) de* A, *il n'existe pas de bijection de* A′ *sur* A. » Un tel ensemble est dit fini et son nombre d'éléments peut se compter. Un ensemble qui ne vérifie pas cette propriété (exemple : $\mathbb{Z}$ et $\mathbb{N} \subset \mathbb{Z}$) est dit infini, et son nombre d'éléments, appelé cardinal infini ou *nombre transfini*, ne peut plus se compter.

OPÉRATIONS SUR LES CARDINAUX

Réunion, intersection d'ensembles ; somme d'entiers

Nous avons déjà remarqué qu'il était indispensable de commettre certains abus de langage ; nous allons en expliciter ici un particulier qui sera commis dans la suite. Lorsque nous dirons : « Étant donné deux objets *a* et *b* », le nombre deux portera sur l'ensemble des signifiants et non sur celui des objets ; nous n'excluons donc pas le cas où les deux symboles désignent le même objet, c'est-à-dire où il n'y a en fait qu'un seul objet. Lorsque nous dirons : « Soient A et B deux ensembles... », nous n'exclurons pas le cas où il n'y a qu'un seul ensemble ◆.

◆ Signalons ici que le mathématicien en disant « deux » pense « au plus deux » et en disant « un » pense « au moins un » !

Sens des mots « ou », « et ». Si E est un ensemble, rappelons qu'un ensemble A est appelé un sous-ensemble de E si tout élément de A est élément de E. Rappelons aussi qu'à toute propriété P(x) définie sur l'ensemble E nous avons associé le sous-ensemble des éléments de E vérifiant la propriété P. Remarquons alors qu'un sous-ensemble A de E est exactement l'ensemble associé à la propriété « $x \in A$ » définie sur E ◆. Une telle propriété est appelée propriété élémentaire ; nous allons, à partir de deux propriétés élémentaires, définir d'autres propriétés sur E, et donc à deux sous-ensembles de E en associer d'autres.

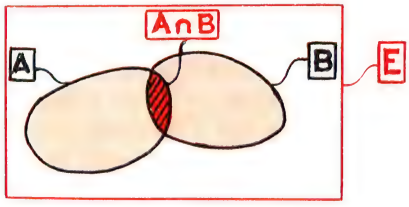
◆ C'est-à-dire que $A = \{ x \in E \mid x \in A \}$

Soient A et B deux sous-ensembles de E ; énonçons les règles nous permettant de définir les propriétés « $x \in A$ et $x \in B$ », puis « $x \in A$ ou $x \in B$ ».

Règle 1. La phrase « $x \in A$ et $x \in B$ » est vraie si les deux phrases « $x \in A$ », « $x \in B$ » sont vraies simultanément et seulement dans ce cas.

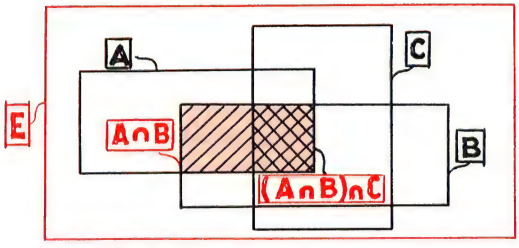
Règle 2. La phrase « $x \in A$ ou $x \in B$ » est vraie si l'une au moins des deux phrases « $x \in A$ », « $x \in B$ » est vraie et seulement dans ce cas.
Les mots « ou », « et » sont pris dans un sens très précis, alors qu'ils peuvent avoir d'autres sens dans le langage courant. Si l'on distingue assez facilement le « et » mathématique du « et » d'énumération, il est plus difficile de comprendre le « ou » mathématique qui est conjonctif, le « ou » du langage ordinaire étant le plus souvent disjonctif (lorsque dans un menu est écrit « fromage ou fruit », le client ne peut avoir les deux aliments à moins de payer un supplément).

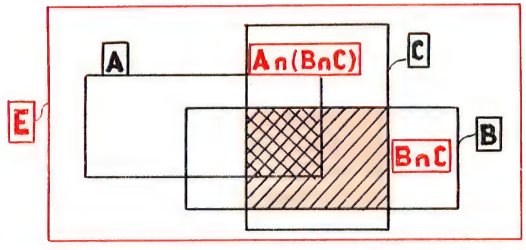
Intersection et réunion. L'intersection de deux sous-ensembles A et B d'un ensemble E est le sous-ensemble de E associé à la propriété « $x \in A$ et $x \in B$ ». On note $A \cap B = \{ x \in E \mid x \in A \text{ et } x \in B \}$ (lire : A inter B).
Par un diagramme de Venn, l'intersection se représente :



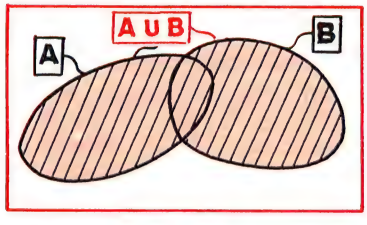
Nous pouvons d'ailleurs définir l'intersection de deux ensembles A et B sans nous placer à priori dans un ensemble. En effet, nous pouvons considérer $A \cap B$ comme le sous-ensemble de A associé à la propriété « $x \in B$ » ; nous aurons $A \cap B = \{ x \in A \mid x \in B \}$ et symétriquement $B \cap A = \{ x \in B \mid x \in A \}$.
Nous avons dit que deux propriétés définies sur le même ensemble sont équivalentes si et seulement si elles définissent le même sous-ensemble. Nous nous contenterons ici de vérifier des égalités entre sous-ensembles à l'aide de diagrammes.

Si A, B, C sont trois sous-ensembles de E, nous notons $A \cap (B \cap C)$ (on lit : A inter ensemble « B inter C ») l'intersection de A et de $B \cap C$, cette intersection étant effectuée précédemment. Constatons l'associativité de l'intersection, c'est-à-dire : $A \cap (B \cap C) = (A \cap B) \cap C$. Nous notons $A \cap B \cap C$ cet unique ensemble.
Nous constaterons de même la commutativité de l'intersection : $A \cap B = B \cap A$. Les égalités suivantes sont encore plus évidentes : $A \cap A = A$ et $A \cap \varnothing = \varnothing$. Enfin, nous dirons de deux ensembles dont l'intersection est vide qu'ils sont disjoints.





La réunion de deux sous-ensembles A et B d'un ensemble E est le sous-ensemble de E associé à la propriété « $x \in A$ ou $x \in B$ ». On note $A \cup B = \{ x \in E \mid x \in A \text{ ou } x \in B \}$ (lire : A union B). Par un diagramme, la réunion se représente :



Si nous voulons définir la réunion de deux ensembles sans nous placer dans un ensemble les contenant, nous sommes obligés d'admettre que la propriété « $x \in A$ ou $x \in B$ » est collectivisante. Nous dressons une liste d'égalités ensemblistes relatives à la réunion en laissant au lecteur le soin de faire les diagrammes correspondants.

$A \cup (B \cup C) = (A \cup B) \cup C$ — *Associativité* de la réunion.
$A \cup B = B \cup A$, — *Commutativité* de la réunion.
$A \cup A = A$ et $A \cup \varnothing = A$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ — *Distributivité de la réunion* par rapport à l'intersection.
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ — *Distributivité de l'intersection* par rapport à la réunion.

Somme de deux nombres

Lorsque nous demandons à un enfant de trouver la somme des nombres deux et trois, il choisira un ensemble modèle de deux, { pouce, index }, puis un ensemble disjoint du précédent modèle de trois, { majeur, annulaire, auriculaire } ; il fera alors la réunion de ces deux ensembles et cherchera le cardinal de la réunion (ici, cinq). *D'une façon générale, si* $x \in \mathbb{N}$ *et* $y \in \mathbb{N}$, *on appelle somme des nombres x et y, notée* $x + y$, *le cardinal de l'ensemble* $A \cup B$, *où* A *et* B *sont deux ensembles disjoints avec* $x = \text{Card } A$ *et* $y = \text{Card } B$. L'opération qui consiste à associer aux deux nombres x et y leur

somme $x + y$ s'appelle addition. Nous n'allons pas, pour chaque addition, manipuler des ensembles d'objets; mais, après, l'avoir fait pour certaines, nous établissons une table d'addition où nous écrivons par exemple $2 + 3 = 5$, table à laquelle nous nous référons pour chaque addition ◆. C'est un acte d'abstraction ; additionner à l'aide de ses doigts, c'est travailler avec des nombres cardinaux, alors qu'utiliser la table c'est opérer avec des symboles abstraits.
Les égalités concernant la réunion se traduisent naturellement en égalités concernant la somme :

De $A \cup B = B \cup A$ on déduit $x + y = y + x$ (commutativité de l'addition).

De $A \cup (B \cup C) = (A \cup B) \cup C$ on déduit $x + (y+z) = (x+y) + z$ (associativité de l'addition).

De $A \cup \varnothing = A$ on déduit $x + 0 = x$.

Nous remarquons que le signe + de l'addition est associé au « ou » ensembliste ; évitons donc de dire « un et un font deux », le « et » utilisé étant celui de l'énumération, non celui du langage mathématique.

◆ Il est à remarquer que si A et A′ sont deux ensembles équipotents de cardinal x : x = Card A = Card A′ et si B et B′ sont deux ensembles équipotents de cardinal y : y = Card B = Card B′ et si A et B sont disjoints ainsi que A′ et B′ alors A ∪ B et A′ ∪ B′ sont équipotents. (Voir définition page 262). Le nombre $x + y$ ne dépend donc que des nombres x et y et non des ensembles choisis pour la manipulation.

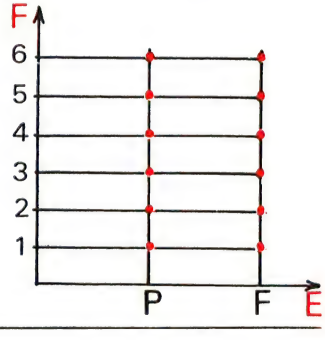
Produits d'ensembles

Couple, ensemble, produit. Si nous jetons une pièce de monnaie, deux résultats sont possibles : soit pile, soit face ; nous avons un ensemble de résultats possibles $E = \{ p, f \}$. De même, si nous jetons un dé, nous avons un ensemble de six résultats possibles $F = \{ 1, 2, 3, 4, 5, 6 \}$. Jetons maintenant en même temps la pièce et le dé, nous convenons de noter le résultat par exemple $(p, 5)$; pour consigner tous les résultats possibles, nous faisons un diagramme cartésien en représentant l'ensemble E horizontalement, F verticalement et marquons les résultats aux nœuds du quadrillage obtenu. Nous avons ainsi construit un nouvel ensemble que nous pourrions écrire en extension :

$\{ (p, 1), (f, 1), (p, 2), (f, 2), \ldots, (p, 6), (f, 6) \}$.

F \ E	P	F
1	(p,1)	(f,1)
2	(p,2)	(f,2)
3	(p,3)	(f,3)
4	(p,4)	(f,4)
5	(p,5)	(f,5)
6	(p,6)	(f,6)

Diagramme cartésien



Nous l'appelons *ensemble produit* de E et de F, noté E × F (lire : E croix F) ; un élément de cet ensemble s'appelle un couple. Nous pouvons également jeter deux fois de suite la même pièce de monnaie et consigner les résultats de manière analogue. Nous obtenons alors l'ensemble produit E × E, noté aussi E^2 ; nous remarquons que le couple (p, f) est distinct du couple (f, p).

E \ E	P	F
P	(p,p)	(f,p)
F	(p,f)	(f,f)

Donnons maintenant des définitions générales. *Si* A *et* B *sont deux ensembles, à tout* $a \in A$ *et à tout* $b \in B$ *nous associons un nouvel objet que nous appelons couple de première composante a et de deuxième composante b, et que nous notons* (a, b). *L'ensemble dont les éléments sont tous les couples s'appelle ensemble produit de* A *et de* B *ou produit cartésien ; il est noté* A × B.

Égalité de deux couples. Les deux écritures (a, b) *et* (a', b') *désignent le même couple, c'est-à-dire* $(a, b) = (a', b')$, *si et seulement si* a *et* a', b *et* b' *désignent les mêmes éléments, c'est-à-dire si* $a = a'$ *et* $b = b'$.

Insistons sur ce dernier point. Si $a \in A$ et $b \in B$, nous pouvons alors considérer l'ensemble $\{a, b\}$ qui s'appelle une paire ; les deux noms paire et couple désignent des objets différents, car, dans le cas où a est distinct de b, les deux paires $\{a, b\}$ et $\{b, a\}$ sont égales comme ayant les mêmes éléments, alors que les couples (a, b) et (b, a) sont distincts ◆.

◆ Pour marquer cette différence, on donne souvent le nom de *paire ordonnée*, voire de *couple ordonné*, à un couple.

Exemple. Dans la vie courante, le mot « paire » peut avoir l'un ou l'autre des deux sens : paire de chaussettes (c'est une paire ; chaque élément est interchangeable), paire de chaussures (c'est un couple). Il en résulte l'inégalité des produits cartésiens A × B et B × A, les deux ensembles A et B étant distincts ; ces deux ensembles produits sont par contre équipotents (considérer la bijection qui à $(a, b) \in A \times B$ associe $(b, a) \in B \times A$).

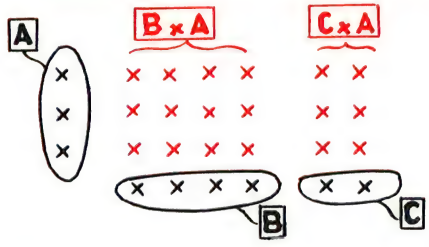
Prenons le singleton $\{0\}$ ◆ ; les deux ensembles A et A × $\{0\}$ sont équipotents (considérer la bijection qui à $a \in A$ associe $(a, 0) \in A \times \{0\}$).

◆ Le singleton $\{0\}$ est l'ensemble formé d'un élément, l'élément 0 (à ne pas confondre avec le 0 qui veut dire nul).

{0} \ A	a	b	c	d	...
0	(a,0)	(b,0)	(c,0)	(d,0)	

D'autre part, pour qu'un couple puisse être défini, c'est-à-dire existe, il faut que les deux composantes existent ; il n'est donc pas possible de construire un couple à partir des éléments de deux ensembles dont l'un est vide ; il en résulte : $A \times \emptyset = \emptyset$.

Nous pouvons généraliser la construction de produits d'ensembles : si A, B, C sont trois ensembles, $(A \times B) \times C$ désigne l'ensemble des couples $((a, b), c)$ et $A \times (B \times C)$, celui des couples $(a, (b, c))$; ces deux ensembles ne sont pas égaux mais équipotents.
Pour terminer, nous donnons une égalité liant réunion et produit cartésien : $(B \cup C) \times A = (B \times A) \cup (C \times A)$; nous la constatons sur le schéma suivant.



Produit de deux nombres. Si $x = \text{Card A}$ *et* $y = \text{Card B}$ *sont deux nombres entiers, on appelle produit de* x *et de* y*, noté* $x \,.\, y$ *ou* $x \times y$*, le cardinal de l'ensemble* $A \times B$.
Les deux entiers x et y s'appellent les facteurs du produit et l'opération qui aux deux facteurs associe le produit s'appelle multiplication.
Nous allons maintenant traduire toutes les propriétés que nous avons énoncées pour le produit cartésien.

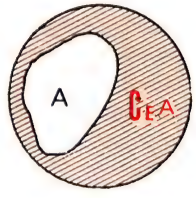
1. L'équipotence de $A \times B$ et de $B \times A$ se traduit : $x.y = y.x$, *commutativité* de la multiplication.
2. L'équipotence de A et de $A \times \{ 0 \}$ se traduit : $x.1 = x$.
3. L'égalité $A \times \varnothing = \varnothing$ se traduit : $x.0 = 0$.
4. L'équipotence de $(A \times B) \times C$ et de $A \times (B \times C)$ se traduit : $(x.y).z = x.(y.z)$, *associativité* de la multiplication.
5. L'égalité $A \times (B \cup C) = (A \times B) \cup (A \times C)$ se traduit : $x.(y + z) = (x.y) + (x.z)$, *distributivité* de la multiplication par rapport à l'addition.

Prenons $A = \{ a, b, c \}$ qui est égal à la réunion des singletons $\{ a \}$, $\{ b \}$ et $\{ c \}$. En utilisant les deuxième et cinquième propriétés, nous avons : $3.x = x + x + x$ ◆. Nous retrouvons ainsi la définition « traditionnelle » : le produit $x.y$ est égal à la somme de x nombres tous égaux à y. Cette dernière définition fait apparaître la multiplication comme une addition répétée, alors que nous l'avons définie de manière indépendante et laissant de plus apparaître la commutativité de façon naturelle.

◆ En effet, soit X un ensemble de cardinal x : $x = \text{Card X}$ et $A = \{ a, b, c \}$ de cardinal 3.
Cherchons à évaluer $\text{Card } (A \times X) = \text{Card A} \times \text{Card X} = 3.x$.
Et, d'après la propriété (5),
$\text{Card } (A \times X) = (\text{Card } \{ a \}).x + (\text{Card } \{ b \}).x + (\text{Card } \{ c \}).x = 1.x + 1.x + 1.x = x + x + x$ d'après (2).

L'ENSEMBLE DES PARTIES D'UN ENSEMBLE

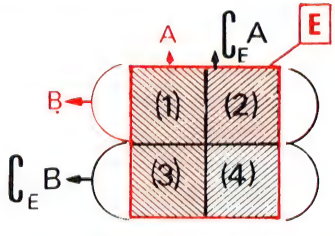
Le mot « non » : complémentaire d'un sous-ensemble



Si E est un ensemble et A un sous-ensemble de E, nous énonçons une nouvelle propriété « non ($x \in A$) », qui obéit à la règle : pour tout élément de E, la phrase « non ($x \in A$) » est vraie si la phrase « $x \in A$ » est fausse et seulement dans ce cas. Dans une initiation au langage ensembliste, sans référence à la logique, il est plus simple de dire « x n'appartient pas à A » et de noter « $x \notin A$ ».

Nous appelons complémentaire de A *dans* E, *noté* $\complement_E A$, *le sous-ensemble associé à la propriété* « *non* ($x \in A$) » ; s'il n'y a pas d'ambiguïté possible sur l'ensemble E, on note simplement $\complement A$. En remarquant que $(\complement A) \cap A = \varnothing$ et $(\complement A) \cup A = E$, nous avons l'égalité Card $(\complement A)$ + Card A = Card E.

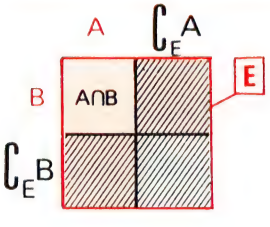
Nous allons constater des égalités, appelées lois de Morgan, liant intersection, réunion et passage au complémentaire à l'aide de schémas appelés diagrammes de Caroll : on partage verticalement le schéma de E (le carré global) en A et $\complement A$ et horizontalement en B et $\complement B$. Nous avons ainsi quatre sous-ensembles disjoints :

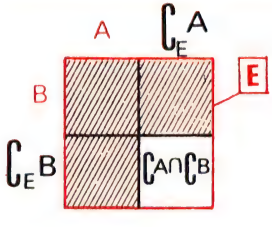


(1) = $A \cap B$
(2) = $B \cap \complement A$
(3) = $A \cap \complement B$
(4) = $\complement A \cap \complement B$

Sur ce schéma, nous remarquons immédiatement que :

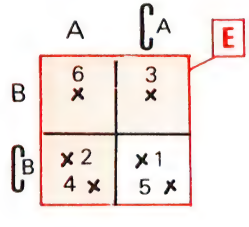
$\complement(A \cap B) = \complement A \cup \complement B$ et $\complement(A \cup B) = \complement A \cap \complement B$ ◆





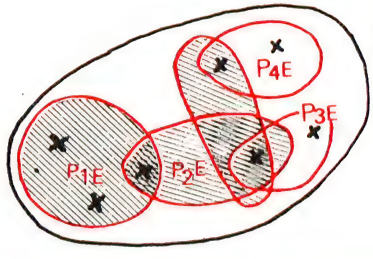
◆ $\complement A \cap B = \complement(1) = (2) \cup (3) \cup (4) = \complement A \cup \complement B$.
$\complement(A \cup B) = (4) = \complement A \cap \complement B$.
La première de ces égalités est liée à l'équivalence des propriétés « non ($x \in A$ et $x \in B$) » et « [non ($x \in A$)] ou [non ($x \in B$)] », la deuxième à l'équivalence des propriétés « non ($x \in A$ ou $x \in B$) » et « [non ($x \in A$)] et [non ($x \in B$)] ».

Par exemple, si E = { 1, 2, 3, 4, 5, 6 }, A = { 2, 4, 6 } et B = { 3, 6 }, nous obtenons le diagramme de Caroll ci-contre.



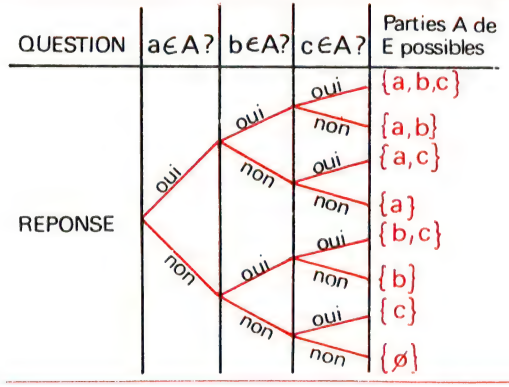
L'ensemble des parties

Nous admettrons que la propriété « être sous-ensemble de l'ensemble E » est collectivisante, c'est-à-dire qu'il existe un ensemble dont les éléments sont les sous-ensembles appelés aussi parties de E ; nous le notons $\mathfrak{P}(E)$. Les phrases « A appartient à $\mathfrak{P}(E)$ » et « A inclus dans E » (notée $A \subset E$) sont donc synonymes. Remarquons que A peut être considéré comme ensemble d'éléments (on écrit alors $x \in A$) ou bien comme élément de $\mathfrak{P}(E)$ (on écrit alors $A \in \mathfrak{P}(E)$). Le même objet peut donc s'écrire à la droite du signe $\in$, dans ce cas c'est un ensemble, ou à la gauche du signe $\in$, dans ce cas c'est un élément. Cela illustre bien le fait qu'on ne peut donner le statut d'élément ou bien celui d'ensemble à un objet, car, selon l'étude qu'on veut faire, il peut être l'un ou l'autre. Jusqu'à présent, nous nous sommes tenus à désigner les éléments par des lettres minuscules et les ensembles par des lettres majuscules ; nous voyons ainsi les limites de la convention que nous avons adoptée jusqu'ici.



P_1E, P_2E, P_3E, P_4E, P_5E... : l'ensemble de tous les regroupements possibles des éléments de E s'appelle $\mathfrak{P}(E)$; P_1E, P_2E, etc. sont à leur tour des éléments de $\mathfrak{P}(E)$.

Arbre des parties. La construction de l'ensemble des parties d'un ensemble peut se faire, dans les cas simples, à l'aide de l'« arbre des parties ».
Donnons un exemple avec $E = \{ a, b, c \}$ et Card $E = 3$.



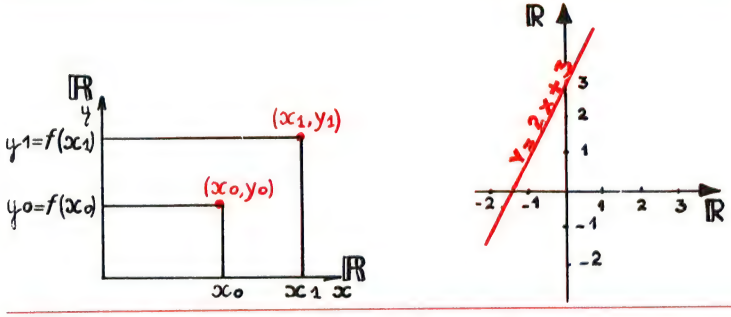
On cherche à construire une partie A de E. On se pose successivement les trois questions : est-ce que $a \in A$, est-ce que $b \in A$, est-ce que $c \in A$? Suivant les réponses à cette question, on obtient toutes les parties A de E possibles.

Soit A un sous-ensemble de E ; pour chaque élément de E, nous nous posons la question : « Cet élément appartient-il à A ? », et nous savons qu'à chaque fois il n'y a que deux réponses possibles : oui ou bien non. En comptant les éléments de $\mathfrak{P}(E)$, nous nous apercevons que leur nombre est huit. Or $8 = 2^3$. Nous donnerons une explication ensembliste à cette égalité.
Dans l'ensemble des parties, nous pouvons définir de nombreuses opérations : intersection, réunion, passage au complémentaire... Nous n'avons donc pas épuisé toutes les possibilités d'opérations sur les ensembles ni donné toutes les propriétés de celles qui ont été déjà définies. Mais nous n'irons pas plus loin, car il n'est pas possible d'étudier ici la structure d'algèbre de Boole de $\mathfrak{P}(E)$.

LES RELATIONS

Dans le second cycle de l'enseignement secondaire, l'étude des « fonctions numériques » (par exemple, $y = ax + b$, $y = \frac{1}{x}$, $y = \sin x$...) a toujours recouvert une grande partie des programmes. Le point de vue adopté était à peu près le suivant : à tout élément x d'un sous-ensemble D de l'ensemble des nombres réels, noté $\mathbb{R}$, une fonction, notée f, fait correspondre un nombre réel y et un seul tel que x et y soient liés par une certaine « expression », notée $y = f(x)$ (par exemple : $y = 2x + 3$). Étudier les « valeurs prises par la fonction », c'était déterminer le sous-ensemble G de $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ des couples ◆ vérifiant l'expression imposée (par exemple : $G = \{ (x, y) \in \mathbb{R}^2 \mid y = 2x + 3 \}$).

◆ Voir plus haut la définition de « couple ».

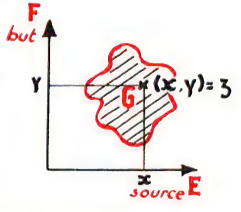


La notion de fonction liée à l'existence d'une expression du type $y = f(x)$ a longtemps été suffisante, mais le cas où à un x donné sont associés plusieurs nombres distincts ne pouvait entrer dans ce cadre. Associons par exemple à tout nombre réel positif les nombres y tels que $x = y^2$. De telles correspondances ne peuvent se décrire à l'aide d'expressions du type $y = f(x)$.

Prenons par exemple $x = 4$. Quel réel allons-nous lui associer tel que $y^2 = 4$? Nous pouvons aussi bien choisir 2 que — 2.
Nous allons voir comment la notion générale de relations permet de décrire toutes ces situations, ainsi que d'autres encore qui n'avaient apparemment aucun point commun avec elles.

Des définitions générales

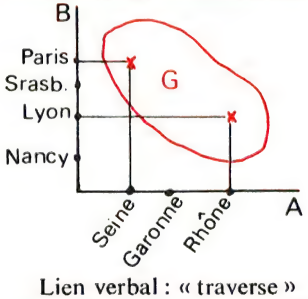
Relations, représentations graphiques. À la fonction liée à l'expression numérique $y = 2x + 3$ nous avons associé un sous-ensemble G de $\mathbb{R}^2$ qui, de ce fait, devait satisfaire à certaines conditions. Pour généraliser, nous partirons d'un sous-ensemble quelconque d'un produit cartésien.
Une relation $\mathcal{R}$ est la donnée d'un produit cartésien $E \times F$ *et d'un sous-ensemble* G *de ce produit cartésien. On la note* : $\mathcal{R} = (E, F, G)$.
Le premier ensemble du produit $E \times F$, c'est-à-dire E, s'appelle la *source* (ou départ), le second, F, le *but* (ou arrivée) et le sous-ensemble G s'appelle le *graphe* de la relation ; si le couple $(x, y) \in G$, x s'appelle l'*antécédent* de y et y l'*image* de x.
Rappelons qu'un sous-ensemble d'un ensemble est souvent déterminé par une *propriété* définie sur cet ensemble. Soit $\mathcal{R} = (E, F, G)$ une relation, le graphe G étant défini par une propriété $P(z)$ sur $E \times F$; si $z \in E \times F$, $G = \{ z \in E \times F \mid P(z) \}$; comme ici z est un couple (x, y), nous noterons cette propriété $x \mathcal{R} y$ et nous aurons $G = \{ (x, y) \in E \times F \mid x \mathcal{R} y \}$ ◆.



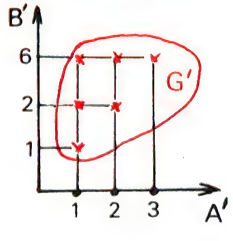
Représentation cartésienne d'une relation $\mathcal{R}$

◆ Lire : G l'ensemble de couples (x, y) appartenant à E croix F telle que x soit en relation avec y.

Donnons deux exemples. Prenons les ensembles $A = \{$ Seine, Garonne, Rhône $\}$, $B = \{$ Paris, Strasbourg, Lyon, Nancy $\}$ et la propriété « x traverse y » définie sur $A \times B$; nous obtenons ainsi la relation (A, B, G) avec $G = \{$ (Seine, Paris), (Rhône, Lyon) $\}$.
Prenons ensuite $A' = \{ 1, 2, 3 \}$, $B' = \{ 1, 2, 6 \}$ et la propriété « x divise y » définie sur $A' \times B'$; nous obtenons ainsi la relation (A', B', G') avec $G' = \{ (1, 1), (1, 2), (1, 6), (2, 2), (2, 6), (3, 6) \}$. Souvent, les ensembles E et F étant donnés, il suffit pour définir la propriété



Lien verbal : « traverse »

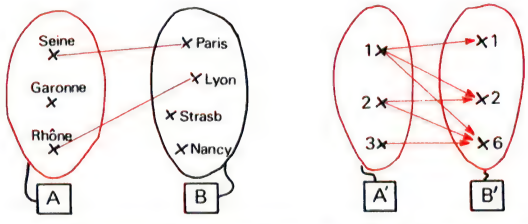


Lien verbal : « divise »

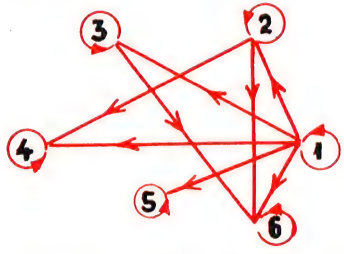
d'un « lien verbal » permettant de « lier » les éléments de E à ceux de F. Dans les exemples précédents, les liens verbaux sont « traverse » et « divise ».
L'étude d'une fonction définie par $y = f(x)$ se terminait par la « représentation graphique ». Ici encore, comme support de l'intuition, nous avons à faire des dessins.

Le schéma cartésien d'une relation consiste à dessiner le produit cartésien, puis à y inscrire les éléments du graphe.

Le schéma sagittal consiste à dessiner les diagrammes de Venn des ensembles départ et arrivée, puis à tracer une flèche de x vers y chaque fois que la phrase « $x \mathcal{R} y$ » est vraie.



Dans le cas où l'ensemble d'arrivée est égal à celui de départ, le schéma cartésien se dessine en traçant deux exemplaires de cet ensemble E ; c'est ce qui a toujours été fait pour les fonctions élémentaires définies par une expression numérique $y = f(x)$, en traçant deux exemplaires de l'ensemble $\mathbb{R}$.
Dans certains cas simples, il est possible de faire un diagramme sagittal en ne prenant qu'un exemplaire de l'ensemble E, puis en traçant les flèches à l'intérieur du diagramme. Ainsi, prenons $E = \{ 1, 2, 3, 4, 5, 6 \}$ et le lien verbal « divise ». Nous obtenons alors le diagramme sagittal ci-dessous :



Égalités de deux relations. Nous dirons que deux relations (E, F, G) et (E′, F′, G′) sont égales si et seulement si E = E′, F = F′, G = G′. Ainsi, nous changeons la relation en changeant la source ou le but, même si l'on conserve le graphe. Reprenons la relation (A′, B′, G′) avec $A' = \{ 1, 2, 3 \}$, $B' = \{ 1, 2, 6 \}$ et

$G' = \{ (x, y) \in A' \times B' \mid x \text{ divise } y \}$; en remplaçant A' par $A'' = \{ 1, 2, 3, 4 \}$, nous ne changeons pas le graphe, mais la relation (A'', B', G') est distincte de la relation (A', B', G').

> Nous n'avons défini en fait que des relations binaires et pourrions également définir une relation n-aire comme la donnée d'un produit d'ensembles $E_1 \times E_2, \ldots, E_n$ et d'un sous-ensemble de ce produit. Cette étude générale n'ayant pas d'intérêt immédiat, nous nous contenterons d'étudier les relations binaires, ce qui nous permet d'ailleurs de ne pas mettre cet adjectif.

Relation réciproque. Étant donné une relation $\mathcal{R} = (E, F, G)$, nous appelons relation réciproque la relation notée $\mathcal{R}^{-1} = (F, E, G^{-1})$ avec $G^{-1} = \{ (y, x) \in F \times E \mid (x, y) \in G \}$; source et but sont donc échangés et le sens des flèches inversé pour le diagramme sagittal. Le passage à la relation réciproque peut se traduire par le passage à la forme passive du lien verbal ; si le lien verbal d'une relation est « divise », celui de la relation réciproque est « est divisé par ».

> Cette définition généralise celle de fonction réciproque ; si, par exemple,
> $\mathcal{R} = (\mathbb{R}, \mathbb{R}, G)$ avec $G = \{ (x, y) \in \mathbb{R}^2 \mid y = 2x + 3 \}$, alors
> $\mathcal{R}^{-1} = (\mathbb{R}, \mathbb{R}, G^{-1})$ avec $G^{-1} = \left\{ (y, x) \in \mathbb{R}^2 \mid x = \dfrac{y - 3}{2} \right\}$.

Relation composée. Prenons deux relations $\mathcal{R} = (E, F, G)$ et $\mathcal{R}' = (F, H, G')$ telles que le but de la première relation soit la source de la seconde ; nous définissons alors la relation composée (E, H, G'') par $G'' = \{ (x, z) \in E \times H$ il existe $y \in F$ avec $(x, y) \in G$ et $(y, z) \in G' \}$.

> Cela généralise la notion de fonction composée ; étant donné deux fonctions dont les expressions sont $z = f(y)$ (par exemple, $z = \sin y$) et $y = g(x)$ (par exemple, $y = 2x + 3$), la fonction composée est la fonction liée à l'expression $z = f(g(x))$ obtenue en remplaçant dans la première expression y par $g(x)$ (par exemple $z = \sin(2x + 3)$).

Utilisant ces notions, nous allons conduire l'étude des relations dans deux directions distinctes :

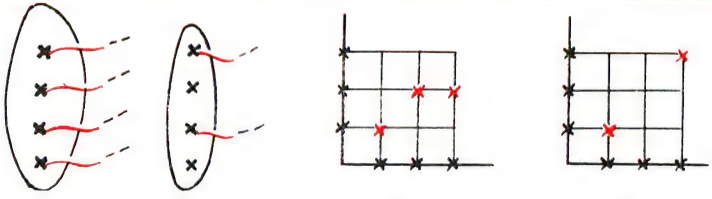
— Étude des relations fonctionnelles d'un ensemble E vers un ensemble F (F pouvant être égal à E), conduisant à la notion d'application et de bijection.

— Étude des relations binaires dans un ensemble E, menant aux relations d'équivalence et d'ordre.

C'est alors que nous pourrons compléter l'étude de l'ensemble $\mathbb{N}$ des nombres entiers.

Relations fonctionnelles

Applications et fonctions. Une relation (E, F, G) est dite fonctionnelle par rapport à sa première variable si pour tout $x \in E$ il existe au plus un $y \in F$ tel que $(x, y) \in G$, ou encore si tout élément de E a au plus une image. Sur un schéma sagittal, on reconnaîtra une telle relation si de tout élément de E part au plus une flèche, et sur un schéma cartésien si une verticale ne coupe le schéma du graphe qu'en un point au plus. Les quatre relations suivantes sont fonctionnelles.



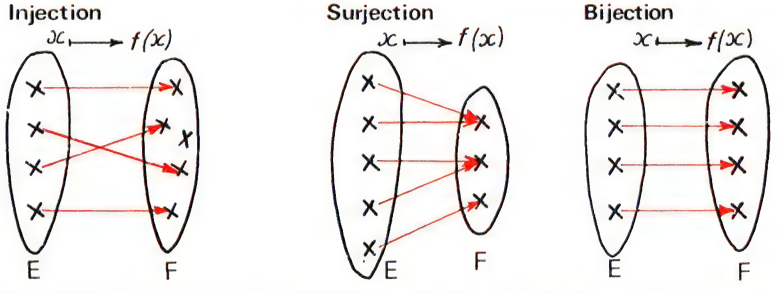
Donnons des exemples plus précis ; la relation $(\mathbb{R}, \mathbb{R}, G_0)$ avec $G_0 = \{ (x, y) \in \mathbb{R}^2 \mid xy = 1 \}$ est fonctionnelle, alors que la relation $(\mathbb{R}, \mathbb{R}, G_1)$ avec $G_1 = \{ (x, y) \in \mathbb{R}^2 \mid y^2 = x \}$ ne l'est pas. De même, la relation $(\mathbb{R}, \mathbb{R}, G_0')$ avec $G_0' = \{ (x, y) \in \mathbb{R}^2 \mid e^y = x \}$ est fonctionnelle, alors que, si $\mathbb{C}$ désigne le corps des nombres complexes, la relation $(\mathbb{C}, \mathbb{C}, G_1')$ avec $G_1' = \{ (x, y) \in \mathbb{C}^2 \mid e^y = x \}$ ne l'est pas.

Dans une relation fonctionnelle, un élément de la source peut ne pas avoir d'image ; *nous réservons le mot « application » aux relations fonctionnelles pour lesquelles tout élément de la source a exactement une image.*

Les relations représentées dans les schémas 1 et 3 sont des applications, alors que les relations représentées dans les schémas 2 et 4 ne le sont pas. Selon les auteurs, le mot « fonction » est synonyme d'« application » ou de « relation fonctionnelle » ; nous réservons ici le mot « fonction » aux applications à valeurs numériques.
Nous noterons :
« $f : E \longrightarrow F$ » une application (E, F, G) ;
« $f(x)$ » l'unique image de $x \in E$;
et « $f(E)$ » le sous-ensemble de F, appelé ensemble image, dont les éléments sont ceux qui ont au moins un antécédent.

Applications injectives, surjectives et bijectives. Si l'application $f : E \longrightarrow F$ est également fonctionnelle par rapport à sa deuxième variable, c'est-à-dire si l'on ne peut avoir $f(x) = f(x')$ que si $x = x'$, alors elle est appelée injective. C'est une *injection.* On peut la définir par une propriété équivalente : $x \neq x'$ entraîne $f(x) \neq f(x')$. Dans

une interprétation sagittale, cela signifie qu'à un élément de F aboutit au plus une flèche.
Si tout élément de F est l'image d'au moins un élément de E, c'est-à-dire si $f(E) = F$, alors l'application est appelée surjective. C'est une *surjection.*
Une application à la fois injective et surjective est bijective et est une *bijection* ; le lecteur reconnaîtra ainsi les bijections introduites de façon intuitive dans la première partie.



La relation réciproque d'une application $f : E \longrightarrow F$ est une application si et seulement si f est bijective ; cette application réciproque est également bijective, nous la notons : $f^{-1} : F \longrightarrow E$.
Si $f : E \longrightarrow F$ et $g : F \longrightarrow H$ sont deux applications, alors la relation composée, f suivi de g, est encore une application que nous notons $g \circ f : E \longrightarrow H$; si les deux applications sont injectives ou bijectives, alors l'application composée l'est également.

Relations dans un ensemble

Propriétés générales des relations. Dans ce paragraphe, nous ne considérons que les relations dont la source est identique au but. Nous commettrons ici un nouvel abus de langage : pour « soit une relation (E, E, G) où G est définie par... », nous dirons « soit dans E une relation définie par... ». Donnons d'abord quelques exemples nous permettant de dégager les principales propriétés.

Exemple 1. Soit un ensemble de six garçons $E = \{ a, b, c, d, e, f \}$; les enfants notés a et c sont frères, ainsi que les enfants notés b, d, f. Considérons dans E la relation définie par « x a les mêmes parents que y » et traçons les diagrammes de cette relation.

Exemple 2. Prenons maintenant $E = \{ 1, 2, 3, 4, 5, 6, 7, 8 \}$ et la relation définie par « x a même reste après division par 3 que y ». Pour trouver les différents éléments du diagramme, il suffit de voir que 2, 5 et 8 ont même reste après division par 3 : reste 2. De même pour 3 et 6 : reste 0, et pour 1, 4 et 7 : reste 1.

Exemple 3. Dans le même ensemble, la relation définie par « $x + y = 10$ ».

Diagramme sagittal

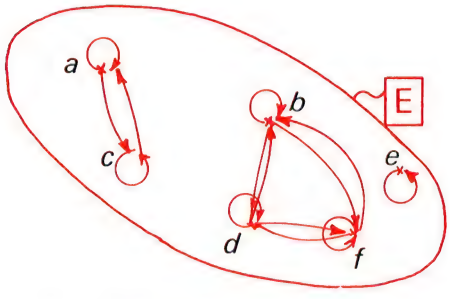
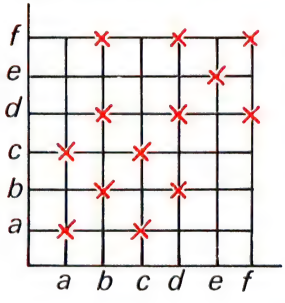


Diagramme cartésien



Exemple 1

Diagramme sagittal

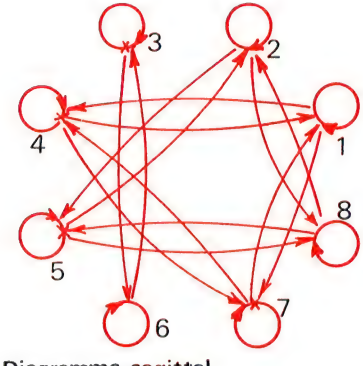
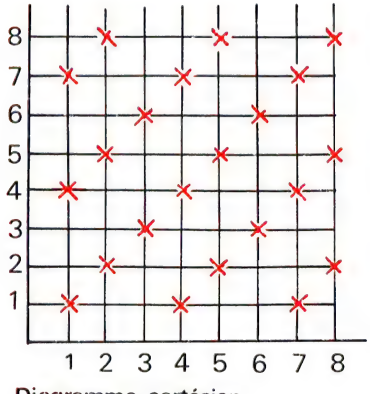


Diagramme cartésien



Exemple 2

Diagramme sagittal

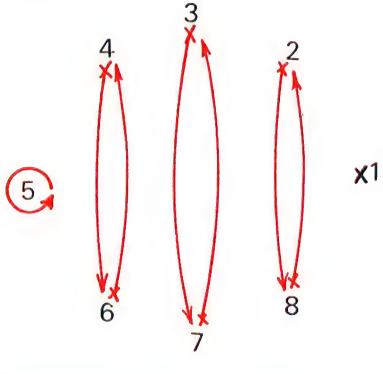
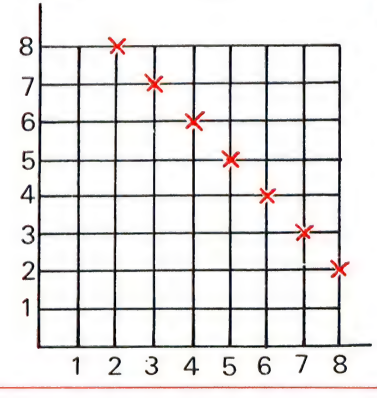


Diagramme cartésien



Exemple 3

Diagramme sagittal

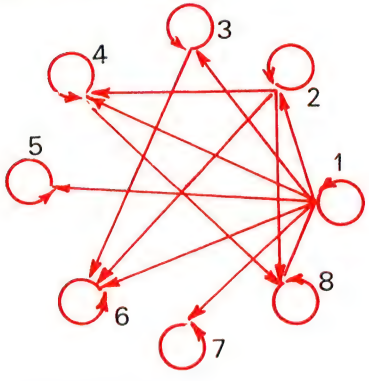


Diagramme cartésien

Exemple 4

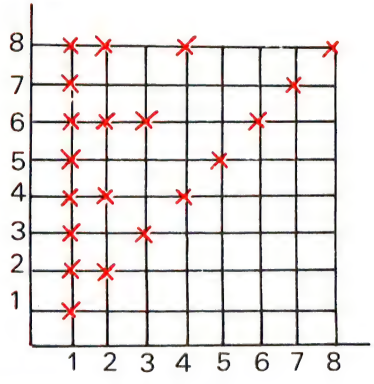


Diagramme sagittal

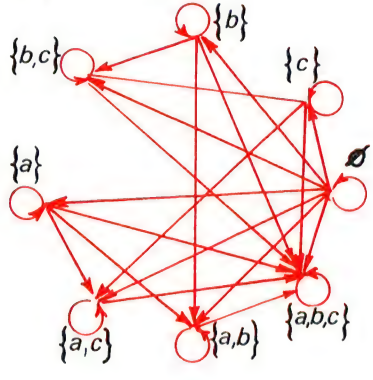
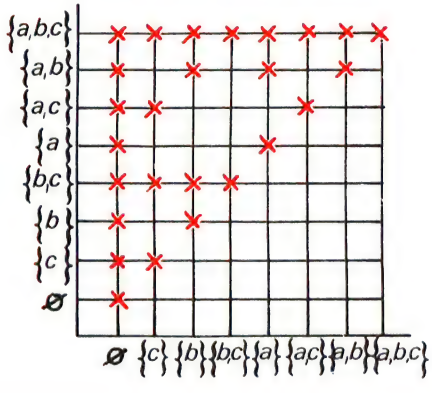


Diagramme cartésien

Exemple 5



Exemple 4. Toujours dans le même ensemble, la relation définie par « x divise y ».

Exemple 5. Soit $E = \{ a, b, c \}$, ensemble dont nous avons dessiné l'arbre des parties ; dans l'ensemble des parties $\mathcal{P}(E)$, considérons la relation définie par « A inclus dans B ».

Réflexivité. Dans les exemples 1, 2, 4 et 5 de chaque point du diagramme sagittal part une flèche qui aboutit en ce point, cela parce que pour tout x la phrase « $x \mathcal{R} x$ » est vraie ; dans ce cas, la diagonale du produit cartésien, ensemble des couples (x, x), est contenue dans le graphe.
On dit qu'une relation dans E de graphe G est réflexive si, pour tout $x \in E$, le couple (x, x) appartient à G.

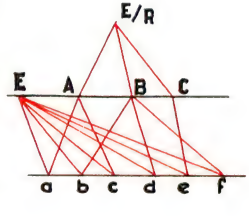
Dans l'exemple 3, le couple (5, 5) appartient au graphe, mais non le couple (2, 2) par exemple, et la relation est non réflexive. Pour démontrer la réflexivité, il faut montrer l'appartenance de tous les couples (x, x) au graphe, alors que pour montrer la non-réflexivité il suffit de montrer qu'il existe un couple du type (x, x) qui n'appartienne pas au graphe.

Symétrie. Dans les exemples 1, 2, 3, chaque fois qu'une flèche part de x et aboutit en y une autre flèche part de y et aboutit en x ; cela parce que pour tout x et tout y si « $x \mathcal{R} y$ » est vraie, alors « $y \mathcal{R} x$ » l'est aussi.
On dit qu'une relation dans E de graphe G est symétrique si, pour tout $x \in E$ et tout $y \in E$, l'appartenance au graphe de (x, y) implique celle de (y, x).
Remarquons que dans l'exemple 1 « $c \mathcal{R} e$ » est faux ; il n'y a donc rien à vérifier sur « $e \mathcal{R} c$ » ; pour vérifier la symétrie, il suffit donc de considérer les couples (x, y) pour lesquels « $x \mathcal{R} y$ » est vraie.

Antisymétrie. Dans les exemples 4 et 5, chaque fois que x et y sont distincts, il n'y a jamais simultanément une flèche de x vers y et une flèche de y vers x ; cela parce que, pour tout x et tout y tels que $x \neq y$, « $x \mathcal{R} y$ » et « $y \mathcal{R} x$ » ne peuvent être vraies simultanément.
Une relation dans E est antisymétrique si pour tout $x \in E$ et tout $y \in E$ tels que $x \neq y$ les couples (x, y) et (y, x) ne peuvent appartenir simultanément au graphe.

Transitivité. Dans les exemples 1, 2, 4, 5, chaque fois que l'on a une flèche de x vers y et une autre de y vers z, il en existe une de x vers z ; cela parce que, si « $x \mathcal{R} y$ et $y \mathcal{R} z$ » est vrai, alors « $x \mathcal{R} z$ » l'est aussi.
Une relation dans E est transitive si, pour tout $x \in E$, tout $y \in E$ et tout $z \in E$, l'appartenance au graphe des couples (x, y) et (y, z) implique celle de (x, z).

Relation d'équivalence. On appelle relation d'équivalence une relation dans un ensemble E qui est à la fois réflexive, symétrique et transitive.
Le sous-ensemble de E dont les éléments sont en relation avec un élément x fixé est appelé *classe d'équivalence* de x, noté Cl (x) ; deux éléments qui sont dans une même classe sont dits équivalents.
Les deux premiers exemples donnent des relations d'équivalence.
Dans le premier exemple, nous avons trois classes :
$A = \text{Cl}(a) = \{ x \in E \mid a \mathcal{R} x \} = \{ a, c \}$, de même $B = \{ b, d, f \}$ et $C = \{ e \}$. Ces trois sous-ensembles sont disjoints deux à deux et leur réunion est tout E ; nous dirons que ces trois parties forment une partition de E. L'ensemble noté E/R dont les éléments sont A, B, C s'appelle *ensemble quotient* ; nous retrouvons une difficulté déjà signalée, à savoir que l'objet A est un élément de E/R, $A \in E/R$, et un ensemble d'éléments, $a \in A$. Le schéma ci-contre permet de



comprendre ce qui se passe ; nous convenons de marquer tous les objets qui interviennent en orientant la page de bas en haut et en reliant par un trait deux points représentant deux objets si celui qui est marqué le plus bas est un élément de celui qui est marqué le plus haut.

Propriété. À toute relation d'équivalence dans un ensemble E nous pouvons associer une partition de E par les classes d'équivalence. D'après la réflexivité, tout élément est dans une classe ; d'après la symétrie et la transitivité, les classes sont disjointes deux à deux : si Cl (x) et Cl (y) ne sont pas disjointes, il existe un élément z dans l'intersection ; les phrases « $x \mathcal{R} z$ », « $y \mathcal{R} z$ » donc « $z \mathcal{R} y$ » sont vraies, et il en est de même de « $x \mathcal{R} y$ » ; par suite, y appartient à Cl (x). Il résulte de la transitivité que Cl (y) est contenue dans Cl (x) ; on montrerait de même l'inclusion de Cl (x) dans Cl (y) et les deux classes sont identiques.

Nous appelons ensemble quotient l'ensemble dont les éléments sont les classes d'équivalence pour une relation d'équivalence donnée ; un élément d'une classe s'appelle un représentant de la classe.
Faisons un bref retour sur la notion d'égalité. Nous avions défini l'égalité comme l'identité. D'un point de vue ensembliste, l'égalité entre éléments n'a de sens que si l'on spécifie l'ensemble dans lequel on se place. Souvent, dans le langage courant, on confond l'égalité des éléments avec l'égalité des classes de ces éléments pour une relation d'équivalence ; lorsque nous disons de deux pièces de monnaie que ce sont les mêmes, nous voulons dire qu'elles sont équivalentes pour la relation « a même valeur que ». Les « segments égaux » et « triangles égaux » donnent une autre illustration de l'abus d'utilisation du mot « égal », car ces segments ou triangles sont seulement équivalents pour la relation « être isométrique à ».

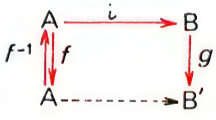
Relation d'ordre. Une relation dans un ensemble qui est à la fois réflexive, antisymétrique et transitive est appelée relation d'ordre ◆. Une relation d'ordre est dite totale si, étant donné deux éléments quelconques x et y, une des phrases « $x \mathcal{R} y$ » ou « $y \mathcal{R} x$ » est vraie. Les deux exemples cités ne sont pas des relations d'ordre total, car 3 ne divise pas 8, 8 ne divise pas 3 ; et $\{ a \}$ n'est pas contenu dans $\{ b, c \}$, $\{ b, c \}$ n'est pas contenu dans $\{ a \}$. Nous verrons plus loin un exemple de relation d'ordre total.

◆ Voir les exemples 4 et 5.

Bon ordre. On dit qu'une relation d'ordre est un bon ordre ou que l'ensemble E est bien ordonné si tout sous-ensemble non vide admet un premier ou plus petit élément. Contre-exemple : l'ensemble $\mathbb{Z}$ des entiers positifs et négatifs n'admet pas de premier élément, car quel que soit l'entier choisi il en existera toujours un avant lui.

Retour sur les cardinaux. Rappelons que nous avons dit que deux ensembles A et B sont équipotents s'il existe une bijection de A dans B. Quelles sont les propriétés de la relation « équipotence »? Nous pouvons remarquer maintenant que la relation « équipotence » est réflexive (l'application Id : E $\longrightarrow$ E définie par $\mathrm{Id}(x) = x$ est une bijection), symétrique (si f : E $\longrightarrow$ F est une bijection, alors la relation réciproque est une bijection de F sur E) et transitive (si f : E $\longrightarrow$ F et g : F $\longrightarrow$ H sont deux bijections, alors la relation composée est une bijection de E sur H). L'équipotence a donc toutes les propriétés d'une relation d'équivalence et n'en est pourtant pas une, car elle n'est pas définie dans un ensemble. Ces remarques justifient la terminologie de classes d'équipotence que nous avons utilisée.

Relation d'ordre dans $\mathbb{N}$. Si $x = \mathrm{Card\ A}$ et $y = \mathrm{Card\ B}$, nous dirons que « x est inférieur ou égal à y », phrase notée « $x \leqslant y$ », si et seulement s'il existe une application injective i de A vers B. Remarquons que, si A est équipotent à A′ (il existe une bijection f: A $\longrightarrow$ A′) et B équipotent à B′ (il existe une bijection g : B $\longrightarrow$ B′), alors on peut construire une application injective de A′ vers B′ (prendre la composée $g \circ i \circ f^{-1}$) ◆ La définition de « x inférieur ou égal à y » ne dépend donc que des classes d'équipotence x et y et non du choix des ensembles A et B dans ces classes. Nous avions la même propriété quand nous avons défini les opérations sur les cardinaux.
Considérons dans l'ensemble des nombres entiers $\mathbb{N}$ la relation définie par « $x \leqslant y$ » ; elle est réflexive (l'application Id : A $\longrightarrow$ A étant injective) et transitive (l'application composée de deux injections est encore injective). Nous admettrons que cette relation est antisymétrique (théorème de Schröder-Bernstein) et que deux cardinaux sont toujours comparables, c'est-à-dire que pour tout A et tout B la phrase « Card A $\leqslant$ Card B ou Card B $\leqslant$ Card A » est vraie (théorème de Zermelo).
La relation dans $\mathbb{N}$ définie par « x inférieur ou égal à y » est donc une relation d'ordre total ; c'est même un bon ordre.



◆ Voir page 20 la composée des applications (dans l'article Algèbre). Rappelons que f^{-1} est l'application réciproque.

Exponentiation de deux cardinaux. Si A et B sont deux ensembles, nous notons A^B l'ensemble des applications de B dans A ; si $x = \mathrm{Card\ A}$ et $y = \mathrm{Card\ B}$, nous définissons x^y (lire x puissance y ou x exposant y) comme le cardinal de A^B. Prenons un exemple : si $A = \{a\}$ et $B = \{b, c\}$, alors $A^B = \{f\}$ avec $f(b) = a$ et $f(c) = a$; par contre, $B^A = \{g, h\}$ avec $g(a) = b$ et $h(a) = c$; nous remarquons donc que $1^2 = 1$ alors que $2^1 = 2$, ce qui montre que l'exponentiation n'est pas une loi commutative ; de plus, en explicitant l'ensemble B^B, nous remarquons que $(2^1)^2 = 2^2$ est distinct de $2^{(1^2)} = 2^1 = 2$; cela montre que la loi n'est pas associative. Nous voyons donc que l'exponentiation est un exemple simple de loi qui ne possède pas les propriétés habituelles de commutativité et d'associativité.

Comme l'ensemble $\varnothing \times \varnothing = \varnothing$ a un seul sous-ensemble, le triplet $(\varnothing, \varnothing, \varnothing)$ est la seule application ayant pour source et pour but l'ensemble vide ; nous aurons Card $(\varnothing^{\varnothing}) = 1$, c'est-à-dire $0^0 = 1$. Si A est un ensemble quelconque, il existe une et une seule application ayant pour source l'ensemble vide et pour but A, à savoir le triplet $(\varnothing, A, \varnothing)$, car $\varnothing \times A = \varnothing$, et nous avons, pour tout cardinal x, $x^0 = 1$. Par contre, si A est non vide, il n'existe aucune application de source A et de but l'ensemble vide ; nous avons, pour $x \neq 0$, $0^x = 0$.

Le théorème de Cantor ◆. Si x et y sont deux cardinaux, nous abrégeons la phrase « x inférieur ou égal à y et x distinct de y » en « x strictement inférieur à y » et nous notons « $x < y$ ».

◆ G. Cantor (1845-1918) : mathématicien allemand qui a consacré sa vie à l'étude de l'infini. Instigateur de la théorie des ensembles, il a offert une vision radicalement nouvelle de la notion de nombre.

Nous voyons intuitivement en faisant l'arbre des parties que, si E est un ensemble fini, alors Card E est strictement inférieur à Card $\mathfrak{P}(E)$; nous allons démontrer que cette inégalité stricte reste vraie si E est un ensemble infini. Ce théorème permet de montrer qu'étant donné un cardinal il en existe un qui soit strictement plus grand. Nous avons Card E $\leqslant$ Card $\mathfrak{P}(E)$; en effet, l'application qui à $x \in E$ associe le singleton $\{ x \} \in \mathfrak{P}(E)$ est une injection de E dans $\mathfrak{P}(E)$. Montrons que l'égalité est impossible ; raisonnons par l'absurde en supposant Card E = Card $\mathfrak{P}(E)$ et en montrant que l'on arrive alors à une contradiction. Supposons qu'il existe une bijection f de E sur $\mathfrak{P}(E)$ et prenons $x \in E$; comme $f(x)$ est un sous-ensemble de E ($f(x) \in \mathfrak{P}(E)$), nous aurons soit $x \in f(x)$, soit $x \notin f(x)$; considérons le sous-ensemble A de E associé à la propriété « $x \notin f(x)$ », $A = \{ x \in E \mid x \notin f(x) \}$. Comme f est surjective, il existe $e \in E$ tel que $A = f(e)$; et nous arrivons alors à une contradiction, car :

— si $e \in A$, il doit vérifier la propriété : $e \notin f(e)$, soit $e \notin A$, ce qui est contradictoire et

— si $e \notin A$, il doit vérifier la propriété : $e \in f(e)$, soit $e \in A$, ce qui est également contradictoire ; l'hypothèse est donc fausse.

Nous allons maintenant exprimer Card $\mathfrak{P}(E)$ en fonction de Card E, plus précisément prouver l'égalité Card $\mathfrak{P}(E) = 2^{\text{Card E}}$. En effet, considérons l'application qui à un sous-ensemble A de E associe sa fonction caractéristique f_A de E dans $\{ 0, 1 \}$ définie par $f_A(x) = 1$ si $x \in A$ et $f_A(x) = 0$ si $x \notin A$.

$$\begin{aligned} f_A : E &\longrightarrow \{ 0, 1 \} \\ x &\longmapsto 0 \text{ si } x \notin A \\ x &\longmapsto 1 \text{ si } x \in A. \end{aligned}$$

À toute partie A de E nous savons associer une application qui va de E dans $\{ 0, 1 \}$ (comme nous venons de le faire). Si A et A′ sont deux parties distinctes, leurs fonctions caractéristiques respectives sont différentes (soit $x_0 \in A$ et $x_0 \notin A'$ alors $f_A(x_0) = 1$ et $f_{A'}(x_0) = 0$).

Inversement, à toute application $f : E \longrightarrow \{ 0, 1 \}$ nous savons associer une partie B de E ainsi définie : $B = \{ x \in E \mid f(x) = 1 \}$. Et à deux applications f_1 et f_2 distinctes nous associerons deux parties B_1 et B_2 distinctes.
Il existe donc une bijection de l'ensemble des parties de E : ensemble $\mathfrak{P}(E)$ dans l'ensemble des applications : $\mathcal{F} : E \longrightarrow \{ 0, 1 \}$. Par conséquent, Card $\mathfrak{P}(E)$ = Card $\mathcal{F}$ = $(\text{Card} \{ 0, 1 \})^{\text{Card E}} = 2^{\text{Card E}}$; en particulier, si Card E est un entier n, alors Card $\mathfrak{P}(E) = 2^n$.

Dénombrement et combinatoire. Nous ne considérons plus que des ensembles finis. Soient E et F deux ensembles ayant respectivement n et p éléments ; nous avons vu que le nombre d'applications de F dans E est n^p ; une telle application s'appelle en combinatoire, lorsque F est un sous-ensemble de E, un arrangement avec répétition de n éléments pris p à p ; il y en a donc n^p. Dénombrons, si $p \leqslant n$, le nombre d'applications injectives de F dans E ; si F est un sous-ensemble de E, une telle application est un arrangement sans répétition de n éléments pris p à p, leur nombre est $A_n^p = n.(n-1).(n-2) \ldots (n-p+1)$. Justifions ce résultat intuitivement : supposons les éléments de F rangés dans un certain ordre $x_1, x_2, \ldots, x_p$ et soit f une injection de F dans E ; pour $f(x_1)$ nous pouvons choisir parmi les n éléments de E, pour $f(x_2)$ parmi les $(n-1)$ éléments de $\complement_E \{ f(x_1) \}$ et pour $f(x_p)$ parmi les $(n-p+1)$ éléments de $\complement_E \{ f(x_1), \ldots, f(x_{p-1}) \}$. Si $n = p$, alors une application injective est bijective et le nombre de bijections de F sur E est $n.(n-1).(n-2) \ldots 3.2.1$, noté $n!$ (se lit : factorielle n) ; en particulier, si E = F, une bijection de E sur lui-même s'appelle une permutation et le nombre des permutations d'un ensemble ayant n éléments est $n!$
Si l'application $f : F \longrightarrow E$ est injective, alors le sous-ensemble image $f(F)$ de E a le même nombre d'éléments que F ; réciproquement, tout sous-ensemble de E ayant p éléments est l'image de F par une application injective de F dans E. Une permutation de F suivie de l'application f est encore une application injective de F dans E dont le sous-ensemble image est encore $f(F)$. Nous avons donc $p!$ applications injectives distinctes de F dans E ayant même sous-ensemble image. Un sous-ensemble de E ayant p éléments s'appelle une combinaison de n éléments pris p à p et leur nombre est : $C_n^p = \dfrac{A_n^p}{p!}$.

Pierre Buisson.

ENSEMBLE D'ARRIVÉE ou BUT : voir Relation

ENSEMBLE DE DÉPART ou SOURCE : voir Relation

ENSEMBLE QUOTIENT : voir Équivalence et la page 259.

ENTIER (Nombre) pages 257 à 273, 264, 265, 323, 326, 366 à 369.

Un nombre entier naturel est un nombre entier positif : 0, 1, 2, 3, ... L'ensemble de ces entiers est noté $\mathbb{N}$.
Un nombre entier positif ou négatif est appelé entier relatif. L'ensemble de ces nombres est noté $\mathbb{Z}$.

$$\mathbb{Z} = \{ \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots \}.$$

Les entiers naturels (entiers positifs) appartiennent à l'ensemble $\mathbb{Z}$: $\mathbb{N} \subset \mathbb{Z}$.

ÉPICYCLOÏDE : voir Cycloïde

ÉPIMORPHISME : voir Homomorphisme

ÉQUATION

On appelle en général équation une relation de la forme $f(x) = 0$, où f est une fonction d'une variable x réelle ou complexe ◆. Il se peut que l'équation ne se présente initialement sous la forme $f(x) = 0$, mais il est toujours possible de l'y ramener.

◆ Variable réelle : les valeurs de la variable sont des nombres réels. Variable complexe : les valeurs de la variable sont des nombres complexes.

Toute valeur x_0 pour laquelle la relation $f(x) = 0$ est vérifiée est appelée solution ou racine de l'équation.

Trouver toutes les racines, c'est résoudre l'équation.

ÉQUATIONS ALGÉBRIQUES pages 194, 327, 333, 440, 441, 475, 521, 522.

Ce sont des équations pour lesquelles la fonction f est une fonction polynôme Le degré du polynôme f est appelé degré de l'équation. Si le degré de l'équation est inférieur à 4, il existe des formules permettant de calculer les racines de l'équation. Nous allons résoudre quelques équations algébriques à coefficients réels et voir ce qu'il convient de changer si les coefficients sont des nombres complexes.

Équation du 1er degré

$$ax + b = 0 \quad (a \neq 0).$$

Il y a une seule racine $x = -\dfrac{b}{a}$ ◆.

◆ La résolution est exactement la même dans $\mathbb{C}$ (ensemble des nombres complexes).

Équation du 2e degré

$$ax^2 + bx + c = 0 \quad (a \neq 0).$$

Cas où les trois coefficients a, b, c sont réels, on commence par calculer le discriminant, noté traditionnellement Δ :

$$\Delta = b^2 - 4ac \text{ ◆}.$$

◆ Nous donnons directement la formule, sans démonstration.

• Si Δ est strictement positif, il y a deux racines réelles x_1 et x_2 données par les formules :

$$x_1 = \frac{-b + \sqrt{\Delta}}{2a} \qquad x_2 = \frac{-b - \sqrt{\Delta}}{2a}.$$

Exemples

$3x^2 + 5x - 2 = 0$

$$\Delta = b^2 - 4ac = (5)^2 - 4 \times 3 \times (-2) = 25 + 24 = 49 > 0.$$

Deux solutions :

$$x_1 = \frac{-5 + \sqrt{49}}{2 \times 3} = \frac{-5 + 7}{6} = \frac{1}{3}$$

$$x_2 = \frac{-5 - \sqrt{49}}{2 \times 3} = \frac{-5 - 7}{6} = -2.$$

• Si Δ est nul, il y a une seule racine, double, $x = -\frac{b}{2a}$.

$3x^2 - 6x + 3 = 0$

$$\Delta = (6)^2 - 4 \times 3 \times 3 = 36 - 36 = 0.$$

Une solution :

$$x_1 = -\frac{-6}{2 \times 3} = \frac{6}{6} = 1.$$

• Si Δ est négatif, il n'y a pas de racines réelles, mais il y a deux racines complexes conjuguées x_1 et x_2 :

$$x_1 = \frac{-b + i\sqrt{-\Delta}}{2a} \qquad x_2 = \frac{-b - i\sqrt{-\Delta}}{2a}.$$

$x^2 + x + 1 = 0$

$$\Delta = 1 - 4 \times 1 \times 1 = -3 < 0.$$

Pas de solution réelle mais deux solutions complexes :

$$x_1 = \frac{-1 + i\sqrt{3}}{2} \qquad x_2 = \frac{-1 - i\sqrt{3}}{2}.$$

Dans le cas où le coefficient b de l'équation est divisible par 2, on emploie, pour résoudre l'équation, le discriminant réduit, noté Δ' : $\Delta' = b'^2 - ac$ (où b' est la moitié de b), et les solutions sont données par les formules :

$$x_1 = \frac{-b' + \sqrt{\Delta'}}{a} \qquad x_2 = \frac{-b' - \sqrt{\Delta'}}{a}.$$

Exemple : $x^2 - 8x + 15 = 0$.

Résolution avec Δ :

$$\Delta = 64 - 4 \times 15 = 4 \quad \begin{cases} x_1 = \dfrac{8+2}{2} = 5 \\ x_2 = \dfrac{8-2}{2} = 3. \end{cases}$$

Résolution avec Δ' :

$$\Delta' = 16 - 15 = 1 \quad \begin{cases} x_1 = \dfrac{4+1}{1} = 5 \\ x_2 = \dfrac{4-1}{1} = 3. \end{cases}$$

Si les coefficients a, b, c sont complexes◆, on calcule aussi Δ (qui, en général, est un nombre complexe), puis les deux racines carrées complexes α et $-\alpha$ de Δ ; il y a deux solutions pour l'équation :

◆ Les coefficients a, b, c sont respectivement égaux aux nombres complexes i, $-(2+i)$, $1-i$.

$$x_1 = \frac{-b+\alpha}{2a} \qquad x_2 = \frac{-b-\alpha}{2a}.$$

$ix^2 - (2+i)x + 1 - i = 0$ i étant un nombre complexe.

$$\Delta = (2+i)^2 - 4i(1-i) = 4 + 4i + i^2 - 4i + 4i^2$$
$$= 4 - 1 - 4 = -1 = i^2$$

Δ a deux racines carrées, i et $-i$; l'équation a donc deux solutions :

$$x_1 = \frac{2+i+i}{2i} = \frac{2+2i}{2i} = \frac{1+i}{i}$$
$$= \frac{(1+i)i}{i^2} = \frac{i+i^2}{-1} = -(i-1) = 1-i$$
$$x_2 = \frac{2+i-i}{2i} = \frac{2}{2i} = \frac{1}{i} = \frac{i}{i^2} = -i.$$

Équation du 3e degré

$ax^3 + bx^2 + cx + d = 0$ (coefficients a, b, c, d réels).

On commence par mettre l'équation sous la forme :

$$X^3 + pX + q = 0$$

en posant $X = ax + \dfrac{b}{3}$. On fait ainsi disparaître le terme du second degré. On calcule le nombre $4p^3 + 27q^2$ (qu'on appelle aussi discriminant de l'équation◆). Si $4p^3 + 27q^2$ est inférieur à 0 (strictement), on a 3 racines réelles distinctes ; si $4p^3 + 27q^2 = 0$, il y a une racine double et une racine simple réelles ; si $4p^3 + 27q^2$ est supérieur à 0, il y a une seule racine réelle et deux autres imaginaires conjuguées.

◆ Et que nous ne démontrons pas.

Voici les formules de Cardan donnant les racines de l'équation dans le cas où $4p^3 + 27q^2$ est positif :

a) Racine réelle

$$X_1 = \sqrt[3]{\alpha} + \sqrt[3]{\beta}$$

b) Racines complexes

$$X_2 = j\sqrt[3]{\alpha} + j^2\sqrt[3]{\beta}$$

$$X_3 = j^2\sqrt[3]{\alpha} + j\sqrt[3]{\beta}$$

où $\alpha = \dfrac{-q + \sqrt{q^2 + \dfrac{4p^3}{27}}}{2} = -\dfrac{q}{2} + \dfrac{1}{2}\sqrt{\dfrac{4p^3 + 27q^2}{27}}$

$$\beta = \frac{-q - \sqrt{q^2 + \dfrac{4p^3}{27}}}{2} = -\frac{q}{2} - \frac{1}{2}\sqrt{\frac{4p^3 + 27q^2}{27}}$$

j, j^2 étant deux nombres complexes dont le cube est égal à 1 :

$$j = \frac{-1 + i\sqrt{3}}{2} \qquad j^2 = \frac{-1 - i\sqrt{3}}{2}.$$

Si $4p^3 + 27q^2 = 0$, il y a une racine simple X_1 et une racine double X_2, qui sont données par les formules :

$$X_1 = \frac{3q}{p} \qquad X_2 = -2X_1.$$

Si $4p^3 + 27q^2$ est négatif, il y a trois racines réelles. Pour résoudre cette équation, on est obligé de faire appel à la trigonométrie. On appelle φ un angle tel que

$$\cos 3\varphi = \frac{3q}{2p}\sqrt{-\frac{3}{p}}.$$

Les solutions sont données par :

$$\begin{cases} X_1 = \sqrt{-\dfrac{4p}{3}}\cos\varphi \\ X_2 = \sqrt{-\dfrac{4p}{3}}\cos\left(\varphi + \dfrac{2\pi}{3}\right) \\ X_3 = \sqrt{-\dfrac{4p}{3}}\cos\left(\varphi + \dfrac{4\pi}{3}\right). \end{cases}$$

Équation du 4e degré

$$ax^4 + bx^3 + cx^2 + dx + e = 0.$$

On met d'abord l'équation sous la forme

$$X^4 + pX^2 + qX + r = 0$$

en posant $X = ax + \dfrac{b}{4}$. On cherche ensuite u de façon que $X^4 + pX^2 + qX + r$ s'écrive, quel que soit X, sous forme du produit

$$(X^2 + uX + v)(X^2 - uX + w)$$

u^2 est racine d'une équation du 3e degré que l'on résout ; la valeur trouvée de u est portée dans l'identité :

$$X^4 + pX^2 + qX + r = (X^2 + uX + v)(X^2 - uX + w)$$

ce qui permet de calculer v et w. Il reste ensuite à résoudre les deux équations du second degré

$$X^2 + uX + v = 0 \text{ et } X^2 - uX + w = 0.$$

Un cas particulièrement simple est le cas de l'équation bicarrée :

$$ax^4 + bx^2 + c = 0$$

que l'on peut résoudre en posant $x^2 = X$. On obtient alors une équation du second degré :

$$aX^2 + bX + c = 0.$$

On résout cette équation et, ensuite, l'équation $x^2 = X$, où l'on a remplacé X successivement par chacune des racines de l'équation $aX^2 + bX + c = 0$.

Si le degré du polynôme est supérieur à 4, pour trouver les racines, on procède par approximations successives.

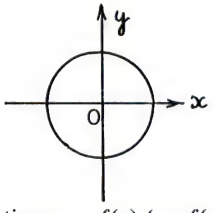
ÉQUATION D'UNE COURBE, D'UNE DROITE, D'UN CERCLE

page 345.

Dans le plan

On appelle équation cartésienne d'une courbe plane la relation de la forme $y = f(x)$ ou $f(x, y) = 0$, liant les coordonnées x et y d'un point M du plan et exprimant une condition nécessaire et suffisante pour que le point M se trouve sur la courbe (tous les points de la courbe ont des coordonnées qui vérifient l'équation et ce sont les seuls).

Exemple



Dans un repère orthonormé

Équation du cercle de centre O et de rayon R :

$$x^2 + y^2 - R^2 = 0.$$

Équation d'une droite non parallèle à Oy :

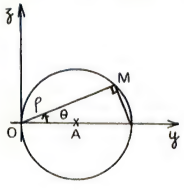
$$y = ax + b.$$

Si la relation $y = f(x)$ (ou $f(x, y) = 0$) est telle que f soit un polynôme, on dit que la courbe est algébrique.

Équation polaire d'une courbe plane

Relation liant les coordonnées polaires ρ et θ d'un point M du plan et exprimant une condition nécessaire et suffisante pour que M appartienne à la courbe. Une telle relation s'écrit :

$$\rho = f(\theta) \text{ ou } f(\rho, \theta) = 0.$$



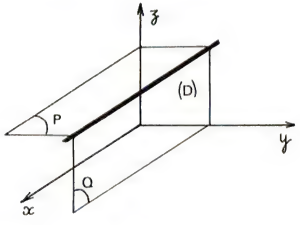
Exemple

Équation polaire du cercle de centre A, de rayon R, et qui passe par l'origine O des axes de coordonnées :

$$\rho = 2R \cos \theta.$$

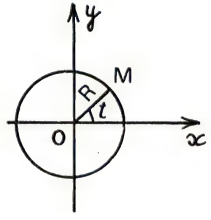
Dans l'espace

Dans l'espace, une courbe est définie en général comme intersection de deux surfaces et elle a donc deux équations.



Par exemple, une droite a deux équations : celles de deux plans dont elle est l'intersection.

Équations paramétriques d'une courbe



Au lieu de traduire le fait que M appartienne à la courbe par une relation liant ses coordonnées x et y, on peut le faire en exprimant chacune de ses coordonnées en fonction d'une même variable (ou paramètre) t : $\begin{cases} x = u(t) \\ y = v(t) \end{cases}$

Ces équations sont les équations paramétriques de la courbe.

Exemple : cercle de centre O et de rayon R : $\begin{cases} x = \mathrm{R} \cos t \\ y = \mathrm{R} \sin t \end{cases}$

Dans l'espace, une courbe aura trois équations paramétriques.

Exemple : hélice circulaire $\begin{cases} x = \mathrm{R} \cos t \\ y = \mathrm{R} \sin t \\ z = h \,.\, t \end{cases}$

ÉQUATIONS AUX DÉRIVÉES PARTIELLES ◆ page 328.

◆ Lire d'abord Fonction, Dérivée et l'article « l'Analyse », pages 125 et suiv.

Équation faisant intervenir plusieurs variables, une fonction de ces variables (qui est l'inconnue de l'équation) et les dérivées partielles de cette fonction, par rapport à chacune de ces variables. Par exemple, une équation aux dérivées partielles du premier ordre, pour une fonction z de deux variables x et y, est une équation de la forme : $\mathrm{F}\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = 0$.

Résoudre cette équation consisterait à trouver la fonction z.

Nous pouvons citer, parmi les équations aux dérivées partielles :

— l'équation de Laplace :

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0$$

(qui fait intervenir les dérivées partielles secondes) ; les solutions de cette équation sont appelées fonctions harmoniques ;

— l'équation des cordes vibrantes :

$$\frac{\partial^2 z}{\partial y^2} - a^2 \frac{\partial^2 z}{\partial x^2} = 0$$

L'étude des équations aux dérivées partielles est encore plus difficile et plus vaste que celle des équations différentielles ; nous ne l'aborderons pas ici.

ÉQUATIONS DIFFÉRENTIELLES ◆ pages 157, 178, 180, 328.

◆ Lire d'abord Équations, Fonction, l'article « l'Analyse » et Différentielle.

Une équation différentielle est une équation faisant intervenir une variable x, une fonction y et les dérivées successives de cette fonction y : y', y'', ..., $y^{(n)}$. Cette équation s'écrit en général $F(x, y, y', y'', \ldots, y^{(n)}) = 0$. Le nombre n est appelé ordre de l'équation différentielle.

Résoudre une équation différentielle, c'est trouver toutes les fonctions y qui satisfont à l'équation.

Une telle fonction est appelée solution de l'équation différentielle ou intégrale de cette équation. En général, ces solutions constituent une famille de fonctions, dont l'équation dépend de un ou de plusieurs paramètres. On appelle ces fonctions intégrales générales de l'équation. Il peut se faire qu'il existe une intégrale qui n'appartienne pas à la famille précédente ; une telle intégrale est appelée intégrale singulière de l'équation. (Par exemple, l'équation différentielle $y' - 2\sqrt{y} = 0$ admet pour intégrales générales les fonctions $y = (x + c)^2$, où c est une constante ; mais la fonction $y = 0$ est aussi intégrale de l'équation sans appartenir à la famille précédente ; c'est une intégrale singulière de l'équation.)

Voici quelques équations différentielles classiques du premier ordre et du second ordre.

Équations du premier ordre

Équation à différentielle totale

C'est une équation qui se présente sous la forme :

$$A(x, y) + B(x, y)\, y' = 0$$

(soit $A(x, y)\, dx + B(x, y)\, dy = 0$, en employant la notation différentielle $y' = \dfrac{dy}{dx}$), où $A(x, y)$ et $B(x, y)$ sont les dérivées partielles, respectivement par rapport à x et par rapport à y, d'une même fonction $u(x, y)$:

$$u'_x . dx + u'_y . dy = 0 \Longrightarrow du = 0.$$

L'intégrale générale est obtenue immédiatement. C'est $u(x, y) = c$.

Équations incomplètes

— Incomplète en y ; c'est une équation de la forme :

$$f(x, y') = 0$$

Exemple : $xy'^2 + xy' - 2y'^2 + x^2 = 0$.

On résout cette équation soit en exprimant y' en fonction de x, soit en paramétrant :

$$\begin{cases} x = u(t) \\ y' = v(t) \end{cases} \Longleftrightarrow \begin{cases} x = u(t) \\ dy = v(t).u'(t)\, dt \end{cases} \Longleftrightarrow \begin{cases} x = u(t) \\ y = \displaystyle\int v(t).u'(t)\, dt. \end{cases}$$

— Incomplète en x ; équation de la forme :

$$f(y, y') = 0$$

Exemple : $y = y'^2 e^{y'}$

que l'on résout de la même façon soit en exprimant y' en fonction de y, soit en trouvant une représentation paramétrique :

$$\begin{cases} y = u(t) \\ y' = v(t) \end{cases} \Longrightarrow \begin{cases} y = u(t) \\ \mathrm{d}x = \dfrac{1}{v(t)} . u'(t)\, \mathrm{d}t \end{cases} \Longrightarrow \begin{cases} y = u(t) \\ x = \displaystyle\int \frac{u'(t)}{v(t)}\, \mathrm{d}t. \end{cases}$$

Équations à variables séparées

Equations de la forme :

$$g(y).y' = f(x)$$

soit encore : $g(y).\mathrm{d}y = f(x).\mathrm{d}x$.

La solution générale est $\displaystyle\int g(y)\, \mathrm{d}y = \int f(x)\, \mathrm{d}x$.

Exemple : $xy' + (y - 1)^2 = 0$

$$x \frac{\mathrm{d}y}{\mathrm{d}x} + (y-1)^2 = 0 \Longleftrightarrow \frac{\mathrm{d}x}{x} = -\frac{\mathrm{d}y}{(y-1)^2} \Longrightarrow y = 1 + \frac{1}{\mathrm{Log}\left|\dfrac{x}{\lambda}\right|}$$

Équations homogènes (en x et y)

Équations de la forme :

$$f\left(y', \frac{y}{x}\right) = 0.$$

Si on peut calculer y' en fonction de $\dfrac{y}{x}$, $y' = \varphi\left(\dfrac{y}{x}\right)$.

On pose $y = tx \Longrightarrow y' = \varphi(t) \Longrightarrow \dfrac{\mathrm{d}y}{\mathrm{d}x} = t + x \dfrac{\mathrm{d}t}{\mathrm{d}x}$.

$$f\left(y', \frac{y}{x}\right) = 0 \Longleftrightarrow \begin{cases} y = t - x \\ x \dfrac{\mathrm{d}t}{\mathrm{d}x} = \varphi(t) - t, \text{ qui est à variables séparées.} \end{cases}$$

Équations linéaires du premier ordre

Équations de la forme :

$$\mathrm{A}(x).y' + \mathrm{B}(x).y = \mathrm{C}(x).$$

On obtient une intégrale générale de cette équation en ajoutant à une intégrale particulière de cette équation une intégrale générale de l'équation sans second membre (c'est-à-dire l'équation $\mathrm{A}(x).y' + \mathrm{B}(x).y = 0$).

Exemple : $y' - y = e^x$.

Une solution particulière de cette équation est $y = x.e^x$ (dont la dérivée est $e^x + x.e^x$).

Les solutions générales de l'équation $y' - y = 0$ sont $y = \lambda.e^x$ (λ est une constante arbitraire).

Donc, les solutions générales de $y' - y = e^x$ sont :

$$y = \lambda.ex + x.ex$$
$$= (\lambda + x)\, e^x.$$

Équation de Bernoulli

Équation de la forme :

$$A(x).y' + B(x).y = C(x).y^n \quad n \neq 1.$$

On la ramène à une équation linéaire en divisant par y^n et en posant $z = \frac{1}{y^{n-1}}$, car l'équation s'écrit alors $\frac{A(x)}{1-n}.z' + B(x).z = C(x)$.

Équation de Ricatti

Équation de la forme :

$$A(x).y' + B(x).y + C(x).y^2 = D(x).$$

On ne sait pas la résoudre dans le cas général. Cela est possible, cependant, si l'on connaît une solution particulière de cette équation. Appelons $u(x)$ cette solution particulière ; si on pose $y = u + z$, l'équation de Ricatti devient une équation de Bernoulli que l'on sait résoudre :

$$y = u + z \Longrightarrow A(x)\,[u' + z'] + B(x)\,[u + z] + C(x)\,[u^2 + 2uz + z^2] = D(x)$$

$Au' + Bu + Cu^2 = D$	$A(x).z' + B(x).z + 2\,C(x).uz + C(x).z^2 = 0$
par hypothèse	$Az' + (B + 2\,Cu).z + Cz^2 = 0.$

Équation de Lagrange et équation de Clairaut

Une équation de Lagrange est une équation de la forme :

$$A(y').x + B(y').y = C(y')$$

(elle est linéaire par rapport à x et y).
Pour la résoudre, on l'écrit :

$$y = \frac{C(y') - x\,A(y')}{B(y')}, \text{ et on pose } y' = z.$$

On différencie $y = \frac{C - xA}{B}$; on y remplace dy par $z\,dx$ (et y' par z).

On obtient alors une équation linéaire en z et x, que l'on sait résoudre. On a aussi x en fonction de z ; on aura y en fonction de z en se servant de la formule $y = \frac{C - xA}{B}$. On obtient donc une représentation paramétrique de la solution.
Cas particulier : *équation de Clairaut*, $y - xy' = f(y')$.
L'intégrale générale est $y = Cx + g(c)$ (équation d'une droite).

Équations du second ordre

La résolution de certaines équations du second ordre peut être faite en se ramenant à une équation du premier ordre. C'est le cas pour les équations incomplètes suivantes :

Équation incomplète en y

Équation de la forme : $f(x, y'\, y'') = 0$.

On pose $y' = z$, donc $y'' = z'$. L'équation devient $f(x, z, z') = 0$, qui est du premier ordre en z. On la résout pour trouver z, puis on résout $y' = z$.

Équation incomplète en x

Équation de la forme : $f(y, y', y'') = 0$.

On pose $y' = z(y)$, d'où $y'' = z' \times z$, ce qui ramène l'équation à une équation du premier ordre en z.

Équations linéaires

Équations de la forme : $A(x)\, y'' + B(x)\, y' + C(x)\, y = D(x)$.

Si on connaît 2 solutions particulières u et v de cette équation, toutes les solutions seront de la forme :

$$\lambda u + \mu v \ (\lambda \text{ et } \mu \text{ constantes}).$$

On obtient l'intégrale générale de cette équation en ajoutant à une intégrale particulière de cette équation l'intégrale générale de l'équation sans second membre.

Cas particulier où $A(x)$, $B(x)$, $C(x)$ *sont constants.* Pour résoudre l'équation sans second membre $Ay'' + By' + Cy = 0$, on cherche, à priori, une solution de la forme $y = e^{rx}$:

$$\Longrightarrow y' = re^{rx} \quad y'' = r^2 e^{rx}.$$

$\Longrightarrow$ L'équation devient $e^{rx}(Ar^2 + Br + C) = 0$, donc ($e^{rx}$ étant toujours différent de 0) $Ar^2 + Br + C = 0$.
Cette équation s'appelle équation caractéristique de l'équation différentielle.
— Si cette équation caractéristique a 2 racines distinctes r_1 et r_2, la solution générale de l'équation $Ay'' + By' + C = 0$ est :

$$y = \lambda e^{r_1 x} + \mu e^{r_2 x}$$

(λ et μ étant 2 constantes arbitraires).
— Si l'équation caractéristique a une racine double r_1, la solution générale de l'équation $Ay'' + By' + Cy = 0$ est :

$$y = e^{r_1 x}(\lambda x + \mu).$$

— Si l'équation caractéristique a deux racines complexes $\alpha + i\beta$ et $\alpha - i\beta$, la solution générale de $Ay'' + By' + Cy = 0$ est :

$$y = e^{\alpha x}(\lambda \cos \beta x + \mu \sin \beta x).$$

Une fois résolue l'équation sans second membre, il faut trouver une intégrale particulière de l'équation avec second membre :
a) Si le second membre $D(x)$ est un polynôme en x, on cherche comme solution particulière un polynôme

- de même degré que $D(x)$ si $C \neq 0$;
- d'un degré supérieur d'une unité au degré de $D(x)$ si $C = 0$ mais $B \neq 0$;
- d'un degré supérieur de 2 unités au degré de $D(x)$ si $C = B = 0$.

b) Si le second membre $D(x)$ est de la forme $e^{mx}.f(x)$, on pose $y = e^{mx}.z$, et l'équation s'écrit :

$$Az'' + (B + 2Am)\, z' + (Am^2 + Bm + C)\, z = f(x).$$

Dans le cas où $f(x)$ est un polynôme, on cherche une intégrale de la forme $e^{mx}.p(x)$ où $p(x)$ est un polynôme de même degré que $f(x)$ si m n'est pas racine de l'équation caractéristique, de degré égal au degré de f plus 1 si m est une racine simple de l'équation caractéristique, de degré égal au degré de f plus 2 si m est racine double de l'équation caractéristique.

c) Si le second membre n'est pas de ces formes, on peut employer la méthode dite de « variation des constantes ». Si u et v sont 2 solutions de l'équation avec second membre, on pose :

$$y = \lambda u + \mu v \ (\lambda \text{ et } \mu \text{ fonctions de } x)$$
$$\Longrightarrow y' = \lambda u' + \mu v' + \lambda' u + \mu' v.$$

On impose $\lambda' u + \mu' v = 0$, donc $y' = \lambda u' + \mu v'$.

Donc $y'' = \lambda' u' + \mu' v' + \lambda u'' + \mu v''$.

$Ay'' + By' + Cy = D$ s'écrit alors :

$$A(\lambda' u' + \mu' v' + \lambda u'' + \mu v'') + B(\lambda u' + \mu v') + C(\lambda u + \mu v) = D.$$

Or $\left.\begin{array}{l} Au'' + Bu' + Cu = 0 \\ Av'' + Bv' + Cv = 0 \end{array}\right\}$ il reste donc $A(\lambda' u' + \mu' v') = D$.

Il reste donc à déterminer λ et μ de façon que :

$$\begin{cases} \lambda' u + \mu' v = 0 \\ \lambda' u' + \mu' v' = \dfrac{D}{A}. \end{cases}$$

ÉQUATION DIOPHANTIENNE

Équation dont on ne recherche de solutions que dans $\mathbb{N}$ (entiers positifs) ou $\mathbb{Z}$ (entiers positifs ou négatifs), c'est-à-dire des solutions entières. Ces équations comportent plusieurs variables. Par exemple :

$$x^2 + y^2 = z^2$$

qui admet pour solutions, entre autres, $x = 3, y = 4, z = 5$ (ou $x = 3k$, $y = 4k$, $z = 5k$, avec k entier quelconque) ou $x = 8$, $y = 6$, $z = 10$ (ou $x = 8k$, $y = 6k$, $z = 10k$).

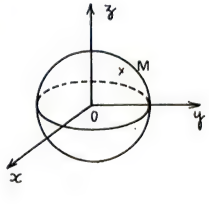
Toutes les solutions de cette équation sont les triplets de nombres (x, y, z) donnés par les formules :

$$x = k(u^2 - v^2) \quad y = 2k.u.v \quad z = k(u^2 + v^2)$$

où u, v, k sont des éléments de $\mathbb{Z}$ (entiers positifs ou négatifs).

ÉQUATION D'UNE SURFACE page 328.

Relation $z = f(x, y)$ ou $f(x, y, z) = 0$, liant les coordonnées d'un point M et exprimant une condition nécessaire et suffisante pour que M appartienne à la surface. Exemples :



— équation de la sphère de centre O et de rayon R :

$$x^2 + y^2 + z^2 - R^2 = 0$$

— équation d'un plan :

$$ax + by + cz + d = 0.$$

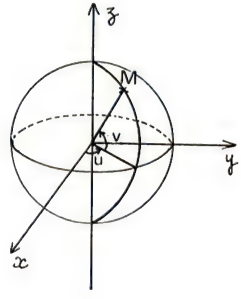
Équations paramétriques d'une surface

Le fait que M appartient à la surface peut se traduire par le fait que ses trois coordonnées sont fonctions de deux paramètres u et v :

$$x = f(u, v) \qquad y = g(u, v) \qquad z = h(u, v).$$

Ces trois équations sont appelées équations paramétriques de la surface. Par exemple, les équations paramétriques de la sphère de centre O et de rayon R sont :

$$x = \mathrm{R} \cos u \cos v \quad y = \mathrm{R} \sin u \cos v \quad z = \mathrm{R} \sin v.$$



ÉQUATIONS TRIGONOMÉTRIQUES page 538.

L'équation la plus simple est l'équation $a \cos x + b \sin x = c$, où a, b, c sont des nombres réels. Pour la résoudre, on peut :

1° *Soit faire un changement de variable* et prendre pour nouvelle variable $\operatorname{tg} \frac{x}{2}$, que l'on pose égale à t. On sait que◆ $\cos x = \frac{1 - t^2}{1 + t^2}$, $\sin x = \frac{2t}{1 + t^2}$ et l'équation s'écrit alors :

◆ Voir Trigonométrie.

$$a \frac{1 - t^2}{1 + t^2} + b \frac{2t}{1 + t^2} = c.$$

Soit encore :

$$(-a - c)\, t^2 + 2bt + a - c = 0 \qquad (1)$$

qui est une équation du second degré en t◆ et que l'on résout. Pour que (1) ait des solutions, il faut que Δ soit positif :

◆ Voir la résolution à Équation.

$$b^2 + (a + c)(a - c) \geqslant 0 \iff b^2 + a^2 - c^2 \geqslant 0 \iff c^2 \leqslant a^2 + b^2.$$

Il faut ensuite, pour avoir x, résoudre les équations $\operatorname{tg} \frac{x}{2} = t$, où t prend successivement les valeurs des solutions de l'équation (1).

2° *Soit faire intervenir le nombre* φ *compris entre* $-\frac{\pi}{2}$ *et* $+\frac{\pi}{2}$, dont la tangente est égale à $\frac{b}{a}$; $\operatorname{tg} \varphi = \frac{b}{a}$ $\left(\text{donc } \cos \varphi = + \sqrt{\frac{a^2}{a^2 + b^2}}\right)$. Mettons a en facteur et remplaçons ensuite $\operatorname{tg} \varphi$ par le rapport $\frac{\sin \varphi}{\cos \varphi}$◆

◆ Voir Trigonométrie.

$$a \cos x + b \sin x = a (\cos x + \operatorname{tg} \varphi \sin x)$$

$$= \frac{a}{\cos \varphi} (\cos \varphi \cos x + \sin \varphi \sin x)$$

$$= \frac{a}{\cos \varphi} \cos (x - \varphi)$$

On a donc à résoudre l'équation $\frac{a}{\cos \varphi} \cos (x - \varphi) = c$, soit :

$$\cos (x - \varphi) = \frac{c \cos \varphi}{a} \left(\cos \varphi \text{ est égal à } + \sqrt{\frac{a^2}{a^2 + b^2}}\right).$$

Pour que cette équation ait des solutions, il faut et il suffit que :

$$-1 \leqslant \frac{c \cos \varphi}{a} \leqslant 1 \Longrightarrow c^2 \leqslant a^2 + b^2.$$

Autres équations trigonométriques simples

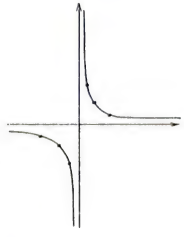
Pour résoudre des équations trigonométriques dont le premier membre est un polynôme en cos x (ou bien un polynôme en sin x ◆), on se ramène à la résolution d'équations algébriques en posant $u = \cos x$ (ou bien $u = \sin x$). Ensuite, on remplace u successivement par toutes les solutions possibles et on résout $\cos x = u$ (ou $\sin x = u$) (seules conviendront les valeurs de u comprises entre -1 et $+1$, puisqu'un cosinus, comme un sinus, est toujours compris entre -1 et $+1$).

◆ Nous nous limiterons aux cas où l'équation ne comporte que des cosinus ou des sinus et non pas les deux. Par exemple, $3\cos^2 x - 2\cos x + 3 = 0$, ou bien $2\sin^2 x - 4\sin x + 1 = 0$.

ÉQUILATÉRAL : voir Polygones

ÉQUILATÈRE (Hyperbole)

Une hyperbole équilatère est une hyperbole dont les deux asymptotes sont perpendiculaires.



ÉQUIPOTENT, ÉQUIPOTENCE ◆ pages 262, 270.

◆ Lire auparavant Bijection.

Soient les deux ensembles $\{1, 2\}$ et $\{a, b\}$; on peut imaginer une bijection du premier vers le second ; par exemple, à 1 on fait correspondre a et à 2 on fait correspondre b ; on dit alors que ces ensembles sont équipotents (voir figure 1). Plus généralement, un ensemble E est équipotent à un ensemble F s'il existe une bijection de l'ensemble E sur l'ensemble F. Cette relation entre ensembles est appelée relation d'équipotence (voir figure 2).

Cette relation est symétrique

Si l'ensemble E est équipotent à F, l'ensemble F est équipotent à E (car il existe une bijection de F sur E : la réciproque de la bijection de E sur F). Au lieu de dire : « E équipotent à F et F équipotent à E », on dit simplement : « E et F sont équipotents » (voir figure 3).

Cette relation est réflexive

L'ensemble E est équipotent à lui-même (voir figure 4).

Cette relation est transitive

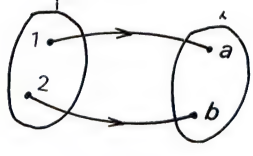
Si, d'une part, E et F sont équipotents et si, d'autre part, F et G sont équipotents, alors E et G sont équipotents.
Si E est un ensemble, on peut rassembler dans une même « classe » tous les ensembles équipotents à E ; on peut associer à cette classe un être mathématique qu'on appelle cardinal de E : Card (E). Cet être est tel que deux ensembles équipotents, donc de la même classe, ont même cardinal, et deux ensembles qui ne

sont pas équipotents, donc de classes différentes, ont des cardinaux différents◆. Le cardinal est caractéristique de la classe d'ensembles à laquelle il a été associé.

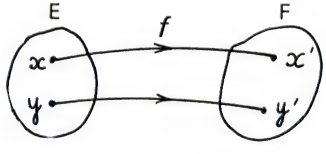
◆ Voir l'article : « les Ensembles ».

Par exemple, à la classe de $E = \{ a, b \}$ est associé le cardinal « deux » ; de façon générale, le cardinal d'un ensemble fini E est le nombre de ses éléments ; le cardinal d'un ensemble infini est un nombre dit transfini ; celui de $\mathbb{N}$ (ensemble des entiers positifs) est noté $\aleph_0$ (aleph zéro), celui de $\mathbb{R}$ (ensemble des nombres réels) est noté $\aleph_1$ (aleph un).
La notion de cardinal est donc une généralisation de la notion intuitive d'entier naturel, mais présente la difficulté suivante : les nombres entiers naturels constituent un ensemble, alors qu'il n'existe pas d'ensemble formé de tous les cardinaux.

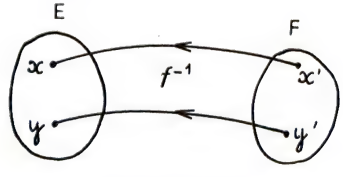
1) Bijection de $\{ 1, 2 \}$ sur $\{ a, b \}$



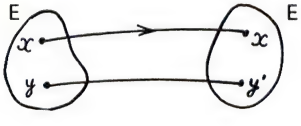
2) Bijection de E sur F



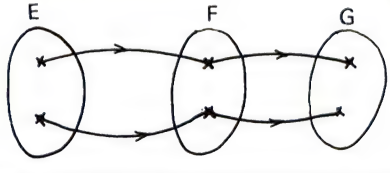
3) Relation symétrique



4) Relation réflexive



5) Relation transitive ◆



◆ Relation d'équipotence : elle est symétrique, réflexive et transitive, mais ce n'est pas une relation d'équivalence.

ÉQUIVALENCE

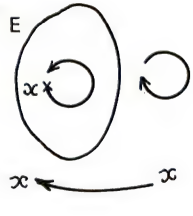
Équivalence logique

Si p et q sont deux propositions, la proposition « $p \iff q$ » (on lit : p est équivalente à q), appelée équivalence logique des deux propositions p et q, est la proposition qui est vraie dans deux cas seulement : le cas où p et q sont vraies toutes les deux et le cas où p et q sont fausses toutes les deux.
Sa table de vérité est représentée ci-contre.

p	q	$p \iff q$
V	V	V
V	F	F
F	V	F
F	F	V

Relation d'équivalence

Une relation $\mathcal{R}$ définie sur un ensemble E est une relation d'équivalence si elle est à la fois réflexive, symétrique et transitive :

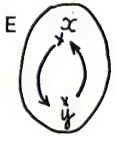


— *Réflexive :*
Chaque élément x de E est en relation avec lui-même :

$$\forall x \in E \qquad x \mathcal{R} x$$

(Lire : quel que soit x appartenant à E, x est en relation avec x).

— *Symétrique :*



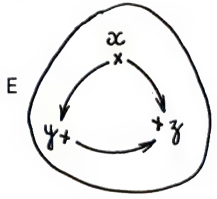
Si x est un élément quelconque de E en relation avec un élément y de E, alors y est en relation avec x :

$$\forall x \in E, \quad \forall y \in E \qquad x \mathcal{R} y \Longrightarrow y \mathcal{R} x$$

— *Transitive :*

Si x est un élément quelconque de E en relation avec y et si y est en relation avec z, alors x est en relation avec z :

$$\forall x \in E, \ \forall y \in E, \ \forall z \in E \qquad x \mathcal{R} y \text{ et } y \mathcal{R} z \Longrightarrow x \mathcal{R} z.$$



Par exemple, l'égalité est une relation d'équivalence sur n'importe quel ensemble E :

$$\forall x \in E \qquad x = x$$

$$\forall x \in E, \ \forall y \in E \qquad x = y \Longrightarrow y = x$$

$$\forall x \in E, \ \forall y \in E, \ \forall z \in E \qquad x = y \text{ et } y = z \Longrightarrow x = z$$

Sur $\mathbb{Z}$ (nombres entiers relatifs négatifs et positifs), la relation : « x a même reste dans la division par n que y » (n étant un entier naturel non nul) est une relation d'équivalence.

Classe d'équivalence

Si E est un ensemble sur lequel est définie une relation d'équivalence $\mathcal{R}$, on appelle classe d'équivalence de l'élément x de E l'ensemble des éléments de E équivalents à x ; on note cet ensemble $\mathrm{Cl}(x)$ (ou $\dot{x}$, ou $\bar{x}$) :

$$\mathrm{Cl}(x) = \{ y \mid y \in E \mid x \mathcal{R} y \}$$

Par exemple, sur $\mathbb{Z}$, soit la relation $\mathcal{R}$ définie par « x a même reste dans la division par 2 que y ».
Dans la division par 2, il y a 2 restes possibles : 0 et 1.

1 divisé par 2 : quotient 0 et reste 1

2 divisé par 2 : quotient 1 et reste 0

3 divisé par 2 : quotient 1 et reste 1
etc.

La classe d'équivalence de 3 contient tous les nombres entiers ayant même reste dans la division par 2 que 3 (c'est-à-dire 1). Cette classe contient donc tous les nombres impairs :

$$\mathrm{Cl}(3) = \{ \ldots, -3, -1, 1, 3, 5, 7, \ldots \} \text{ (nombres impairs).}$$

La classe d'équivalence de 2 contient tous les nombres dont le reste dans la division par 2 est le même que celui de 2, c'est-à-dire 0 ; elle contient donc tous les nombres pairs :

$$\mathrm{Cl}(2) = \{ \ldots, -2, 0, 2, 4, 6, \ldots \} \text{ (nombres pairs).}$$

Une classe d'équivalence $\mathrm{Cl}(x)$ n'est jamais vide, car elle contient toujours x.

Si a et b sont deux éléments de E, leurs classes d'équivalence sont soit égales si a est équivalent à b, soit disjointes si a est non équivalent à b.
Les classes d'équivalence de tous les éléments d'un ensemble E (pour la relation $\mathcal{R}$) constituent une partition de E.

Ensemble quotient

L'ensemble des classes d'équivalence de E est appelé ensemble quotient de E par $\mathcal{R}$ et noté E/$\mathcal{R}$.
Si E = $\mathbb{Z}$ et la relation $\mathcal{R}$: « x a même reste dans la division par 2 que y », l'ensemble quotient a deux éléments : Cl(0) (nombres pairs) et Cl(1) (nombres impairs). E/$\mathcal{R}$ est ici noté $\mathbb{Z}/2$ et $\mathbb{Z}/2 = \{ \text{Cl}(0), \text{Cl}(1) \}$.

ÉQUIVALENT ◆ pages 260, 523.

◆ Lire d'abord Equivalence (relation d').

Soit E un ensemble sur lequel existe une relation d'équivalence $\mathcal{R}$; si x est un élément de E en relation avec l'élément y de E, on dit que x est équivalent à y ; comme la relation $\mathcal{R}$ est une relation symétrique, y est aussi équivalent à x ; on dit donc simplement que x et y sont équivalents.

ERREUR

En calcul numérique, on appelle erreur la différence entre la valeur calculée d'un nombre et sa valeur exacte. Cette erreur peut être positive ou négative selon que le calcul a été fait par excès ou par défaut. La valeur exacte d'une erreur est, naturellement, inconnue et souvent, même, son signe l'est aussi ; ce que l'on peut déterminer, c'est un majorant de la valeur absolue de cette erreur : c'est ce qu'on appelle l'incertitude.

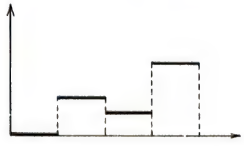
Par exemple, si on prend pour valeur de π : 3,14, l'erreur commise est 3,14 — π. Une valeur approchée plus précise de π étant 3,141 6, l'incertitude peut être prise égale à 0,002 (qui est bien supérieure à la valeur absolue de la différence 3,14 — 3,141 6).

ESCALIER (Fonction en)◆

◆ Lire d'abord Intervalle.

Une fonction réelle d'une variable réelle, définie sur un intervalle (a, b), est une fonction en escalier si on peut partager (a, b) en intervalles disjoints (a_1, b_1), (a_2, b_2), ... tels que, sur chacun d'eux, la fonction soit constante.

Représentation graphique d'une fonction en escalier



ESPACE

Espace affine : voir Affine et Algèbre linéaire, page 62.

Espace homogène : voir l'article « l'Algèbre », page 45.

Espace topologique : voir Topologie

Espace vectoriel : voir Vectoriel (espace) et Algèbre linéaire, page 55.

ESPÉRANCE MATHÉMATIQUE : voir l'article « les Probabilités », page 428, l'article « les Statistiques », page 485 et la page 235.

ET (Conjonction) pages 98, 265, 310, 416, 364.

Si p et q sont deux propositions, la proposition « p et q » est la proposition qui est vraie si et seulement si les deux propositions p et q le sont toutes les deux. Cette proposition est aussi appelée conjonction des deux propositions p et q. Sa table de vérité est représentée ci-contre.

p	q	p et q
V	V	V
V	F	F
F	V	F
F	F	F

EUCLIDIEN : voir Distance et Norme

EULER (relation d') : voir Polyèdres

ÉVÉNEMENT : voir Probabilités, page 414 et suivantes.

EXPONENTIATION, EXPOSANT, PUISSANCE pages 22, 25, 327.

Exposant entier positif

Le carré d'un réel x est égal au produit de x par lui-même ; on le note x^2 (on lit : x deux, ou x exposant deux, ou x au carré) ; $x^2 = x \times x$.
Le cube de x, noté x^3, est égal au produit de x^2 par x :

$$x^3 = x^2 \times x = (x \times x) \times x = x \times (x \times x) = x \times x^2 = x \times x \times x.$$

De façon générale, si n est un entier naturel positif supérieur à 1, on note x^n le produit de x^{n-1} par x, c'est-à-dire le produit de n termes égaux à x :

$$x^n = x^{n-1} \times x = \underbrace{x \times x \times x \times \ldots \times x}_{n \text{ fois}}$$

Pour $n = 1$, on pose $x^1 = x$◆.
Quels que soient les nombres réels x et y et quels que soient les entiers m et n (non nuls), les propriétés suivantes sont vraies :

$$\text{(F)} \begin{cases} (x.y)^n = x^n.y^n \\ x^m.x^n = x^{m+n} \\ (x^m)^n = x^{m \cdot n} \end{cases}$$

◆ Dans le cas de $\mathbb{N}$ (entiers naturels positifs), pour tout entier naturel n ($n \neq 0$) et tout entier naturel x, x^n est un entier naturel. On définit ainsi une loi de composition interne sur $\mathbb{N}$, qu'on appelle exponentiation. Cette loi n'est ni commutative ($2^3 \neq 3^2$) ni associative ($2^{(1^2)} \neq (2^1)^2$).

Exposant entier négatif

Chaque élément x non nul de $\mathbb{R}$ admet un inverse $\frac{1}{x}$; on définit alors :

$$x^{-1} = \frac{1}{x} \text{ (« } x \text{ puissance — 1 »)}$$

$$x^{-2} = \frac{1}{x} \times \frac{1}{x} = \frac{1}{x^2} \text{ (« } x \text{ puissance — 2 »)}$$

et, de façon générale, « x puissance — n » (n entier naturel $\neq 0$) comme étant le produit de n termes égaux à $\frac{1}{x}$, ce qui est encore égal à l'inverse de x^n :

$$x^{-n} = \underbrace{\frac{1}{x} \times \frac{1}{x} \times \ldots \times \frac{1}{x}}_{n \text{ fois}} = \frac{1}{x^n}$$

Les exposants ainsi définis sont des entiers négatifs pour lesquels les formules (F) sont encore valables.

Si x est un réel non nul : $x^m \times \frac{1}{x^n} = x^m \times x^{-n} = x^{m-n}$

En particulier, si $m = n$: $x^m \times \frac{1}{x^m} = 1 = x^0$

Pour tout réel $x \neq 0$, $x^0 = 1$.

Exposant rationnel

Si q est un entier naturel supérieur à 1 et si x est un nombre réel strictement positif, la racine q-ième de x existe et est positive (c'est le nombre positif y qui, élevé à la puissance p, est égal à x) ; quel que soit l'entier relatif p, on définit alors $x^{\frac{p}{q}}$ comme étant la racine q-ième de x élevée à la puissance p :

$$x^{\frac{p}{q}} = \left(\sqrt[q]{x}\right)^p \blacklozenge$$

Les propriétés F sont encore valables pour ces exposants rationnels, à condition que les réels x et y qui figurent dans F soient strictement positifs.

◆ $y = \sqrt[q]{x} \iff x = y^q$; $x^{\frac{1}{2}} = \sqrt{x}$; $x^{\frac{2}{3}} = \left(\sqrt[3]{x}\right)^2$; $x^{-\frac{1}{2}} = \frac{1}{\sqrt{x}}$.

Puissance d'un ensemble

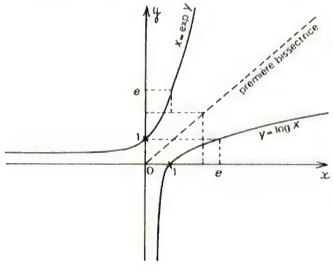
Dire : « Les deux ensembles E et F sont équipotents » est synonyme de : « Les deux ensembles E et F ont même puissance. »

EXPONENTIELLE (Fonction) ◆ page 159, 283.

La fonction exponentielle est la fonction réciproque de la fonction logarithme népérien. Cette dernière (notée Log) est une application de l'intervalle ouvert $]0, +\infty[$ dans l'ensemble des nombres réels $\mathbb{R}$, continue et strictement croissante ; elle admet une application réciproque ◆, qui est une application de $\mathbb{R}$ dans $]0, +\infty[$, continue et strictement croissante : on l'appelle exponentielle (et on la note exp). Si à x correspond y par la fonction Log ($y = \text{Log } x$), à y correspond x par la fonction exp ($x = \exp y$) : $y = \text{Log } x \iff x = \exp y$.
La courbe représentative de la fonction exponentielle (dans un repère orthonormé) est symétrique de celle de la fonction Log par rapport à la première bissectrice.

◆ Lire d'abord Logarithme et Réciproque (fonction).

◆ Voir dans l'article « l'Analyse », page 122.



Quel que soit le nombre entier relatif n, on démontre que exp n est égal à e^n (e étant le seul nombre pour lequel le logarithme népérien vaut 1 ; $e = 2{,}718\,28...$). On convient d'écrire, pour tout nombre réel x, exp $x = e^x$. On dit que la fonction exponentielle est la fonction exponentielle de base e.

EXTENSION

Un ensemble E est défini en extension s'il est défini par la liste de tous ses éléments. Par exemple, l'ensemble des voyelles de l'alphabet français est défini en extension lorsqu'on donne la liste des 6 voyelles a, e, i, o, u, y ; cet ensemble est noté : $\{ a, e, i, o, u, y \}$ en faisant figurer entre accolades la liste de ses éléments (séparés par une virgule, et chacun d'eux n'apparaissant qu'une fois).

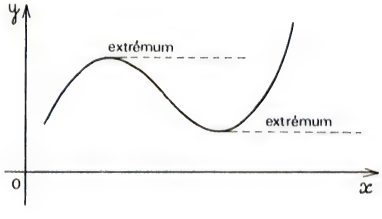
EXTÉRIEUR : voir Intérieur

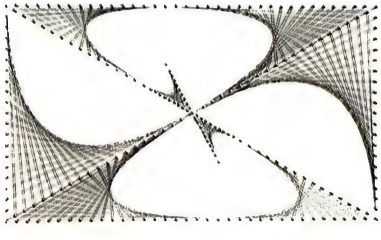
EXTERNE (Loi de composition) : voir Composition (loi de)

et les pages 60, 61.

EXTRÉMUM (d'une fonction) page 311.

Mot employé pour indiquer qu'une fonction f d'une variable réelle x admet un maximum ou un minimum, sans préciser s'il s'agit de l'un ou de l'autre. S'il y a extrémum pour une valeur x_0 de la variable, alors la dérivée de la fonction s'annule pour cette valeur x_0. La réciproque n'est vraie que si la dérivée s'annule *en changeant de signe*.





F

FACE : voir Polyèdres

FACTEUR

Un facteur est un des termes d'un produit. Par exemple, 2×3 comporte deux facteurs : 2 et 3.

Mettre un terme en facteur dans une expression algébrique, c'est utiliser la distributivité de la multiplication par rapport à l'addition ◆, pour écrire l'expression sous la forme d'un produit. Par exemple, dans l'expression :

$$(x - 2)(2x + 3) - (x - 2)(x + 1),$$

on peut mettre $(x - 2)$ en facteur :

$$(x - 2)[(2x + 3) - (x + 1)] = (x - 2)(x + 2).$$

◆ Pour tous nombres réels a, b, c : $(a \times b) + (a \times c) = a \times (b + c)$.

FACTORIELLE pages 35, 285.

Si n est un nombre entier naturel (positif), on appelle factorielle n, et on note $n!$ (lire : factorielle n), le produit de tous les entiers consécutifs depuis 1 jusqu'à n. Par exemple :

$$3! = 1 \times 2 \times 3$$
$$5! = 1 \times 2 \times 3 \times 4 \times 5.$$

0! n'a pas d'existence selon la définition ci-dessus ; on a été amené à poser par convention $0! = 1$ pour que la formule $\frac{n!}{p!\,(n-p)!} = 1$ ◆ soit vraie si $n = p$ $[(n - p)! = 0! = 1]$.

◆ Voir Combinaison.

FAISCEAU page 208.

Si (C) et (C′) sont deux courbes planes d'équations respectives $f(x, y) = 0$ et $g(x, y) = 0$, on appelle faisceau linéaire engendré par (C) et (C′) l'ensemble de toutes les courbes du plan dont l'équation est de la forme :

$$f(x, y) + \lambda g(x, y) = 0,$$

où λ est un nombre réel quelconque.

Si les deux courbes (C) et (C′) se coupent en un certain nombre de points, *toutes* les courbes du faisceau passent par ces points.

Faisceau de droites

Un faisceau de droites est constitué soit par des droites toutes parallèles entre elles (si (C) et (C′) sont deux droites parallèles), soit par des droites passant toutes par le même point (si (C) et (C′) sont deux droites concourantes).

Faisceau de cercles

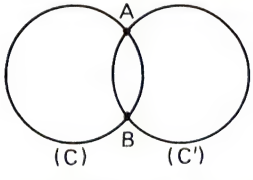
Il y a trois types de faisceau de cercles :
— Si les deux cercles (C) et (C′) se coupent en A et B, tous les cercles du faisceau passent par A et B ; A et B sont appelés points de base du faisceau.

— Si (C) et (C′) sont tangents en A, tous les cercles du faisceau passent par A et sont tangents en A.

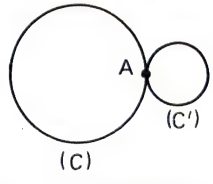
— Si (C) et (C′) ne se coupent pas, sur la droite (D) joignant les centres des deux cercles il existe deux points I et J tels que les divisions A, B, I, J et A′, B′, I, J soient harmoniques ◆. Chaque cercle du faisceau coupe la droite (D) en deux points tels que la division formée par ces deux points et I et J soit harmonique. I et J sont appelés points limites du faisceau.

◆ Voir Birapport.

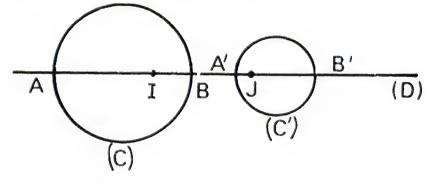
Faisceau à points de base



Faisceau de cercles tangents



Faisceau à points limites



FAISCEAU HARMONIQUE : voir Birapport

FERMÉ ◆ page 13.

◆ Vocabulaire de topologie. Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

Si E est un espace topologique, on appelle partie fermée de E, ou, plus simplement, fermé de E, tout sous-ensemble de E dont le complémentaire dans E soit un ouvert de E. Soit $\mathbb{R}$ (ensemble des nombres réels) muni de la topologie de l'ordre ; les intervalles fermés de $\mathbb{R}$ $\left([a, b] = \left\{ x, a \leqslant x \leqslant b \right\}\right)$ sont des fermés de E (puisque complémentaire de l'ouvert $]-\infty, a[\cup]b, +\infty[$).
Si E est muni de la topologie grossière, les seuls fermés de E sont l'ensemble vide ∅ et l'ensemble E (qui se trouvent donc être à la fois ouverts et fermés). Si E est muni de la topologie discrète, tout sous-ensemble de E est à la fois ouvert et fermé.
Les propriétés des fermés se déduisent immédiatement de celles des ouverts :

a) toute intersection de fermés est un fermé ;

b) toute réunion finie de fermés est un fermé ;

c) E et ∅ sont des fermés.

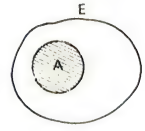
Comme il est équivalent de se donner le système d'ouverts de E ou le système de fermés, on pourrait définir une topologie sur un ensemble E à partir d'une famille $\mathcal{F}$ de fermés.

FILTRE ◆

◆ Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire l'article « les Ensembles ».

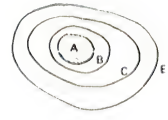
Soient E un ensemble et A une partie non vide de cet ensemble. Appelons $\mathcal{F}$ la famille de toutes les parties de E qui contiennent A. Cette famille $\mathcal{F}$ n'est pas vide (il y a au moins E qui en fait partie) ; l'ensemble vide ∅ n'appartient pas à cette famille (l'ensemble vide ne peut contenir un ensemble non vide) ; si B est un

élément de cette famille et si C est une partie de E qui contient B, alors C appartient à la famille $\mathcal{F}$ (puisque C contient B qui contient A, C contient donc A). Si B et C sont deux éléments de la famille $\mathcal{F}$, tous deux contiennent A ; leur intersection contient donc aussi A ; elle appartient à la famille $\mathcal{F}$.
On dit que cette famille $\mathcal{F}$ est un filtre sur l'ensemble E.
De façon générale, un filtre $\mathcal{F}$ sur un ensemble E est un ensemble de parties de E qui satisfait aux conditions suivantes :



1) $\mathcal{F}$ n'est pas vide et l'ensemble vide n'appartient pas à $\mathcal{F}$:

$$\mathcal{F} \neq \varnothing \qquad \varnothing \notin \mathcal{F}$$

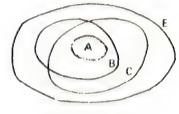


2) Si F appartient à $\mathcal{F}$, alors toute partie G (de E) contenant F appartient à $\mathcal{F}$:

$$F \in \mathcal{F} \qquad F \subset G \Longrightarrow G \in \mathcal{F}$$

3) Si F appartient à $\mathcal{F}$ et si F 'appartient à $\mathcal{F}$, alors l'intersection de F et de F' appartient à $\mathcal{F}$:

$$\left.\begin{array}{l} F \in \mathcal{F} \\ F' \in \mathcal{F} \end{array}\right\} \Longrightarrow F \cap F' \in \mathcal{F}$$



Autres exemples

a) L'ensemble des voisinages d'un point x d'un espace topologique E est un filtre appelé filtre des voisinages de x.

b) Dans $\mathbb{N}$ (ensemble des entiers naturels positifs), l'ensemble constitué par les complémentaires des parties finies de $\mathbb{N}$ est un filtre appelé filtre de Fréchet ◆.

◆ Maurice Fréchet : mathématicien français contemporain.

FINE (Topologie plus) ◆

Soit E un ensemble que l'on peut munir de deux topologies, l'une dont la famille d'ouverts est notée $\mathcal{O}_1$ et l'autre dont la famille d'ouverts est notée $\mathcal{O}_2$; la topologie définie par $\mathcal{O}_1$ est moins fine que la topologie définie par $\mathcal{O}_2$ si tout élément de $\mathcal{O}_1$ est élément de $\mathcal{O}_2$.

◆ Vocabulaire de topologie. Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

La topologie grossière est la moins fine de toutes les topologies que l'on peut définir sur un ensemble E, car les deux ouverts de cette topologie : l'ensemble vide $\varnothing$ et E, sont ouverts de E pour n'importe quelle autre topologie.

La topologie discrète est la plus fine de toutes les topologies ◆ que l'on peut définir sur un ensemble E, car tout ouvert de n'importe quelle topologie sur E est une partie de E, donc un ouvert de E pour la topologie discrète (pour laquelle les ouverts de E sont les parties de E).

◆ C'est pour cette raison qu'on l'appelle aussi parfois topologie fine.

La relation ainsi définie entre deux topologies sur E est une relation d'ordre.
Si la topologie définie par $\mathcal{O}_1$ est moins fine que la topologie définie par $\mathcal{O}_2$, alors l'application identique (qui à x de E fait correspondre x de E) de E muni de la topologie définie par $\mathcal{O}_2$ dans E muni de la topologie définie par $\mathcal{O}_1$ est continue, et réciproquement (l'application identique de E muni de la topologie $\mathcal{O}_1$ dans E muni de la topologie $\mathcal{O}_2$ n'est pas continue).

FINI (Ensemble) : voir l'article « les Ensembles », page 264 et l'article « les Nombres », page 375 et les pages 105, 265, 299, 375, 377, 418.

FLOU (Mathématique du)

La rigueur mathématique tient au caractère extrêmement simple de son mode de pensée. La démarche, élémentaire, consiste à poser deux possibilités qui, s'excluant l'une l'autre, forment alternative, en sorte que le point est dans le plan ou sinon en dehors ; il est ou non le centre du cercle. Les deux droites sont ou non perpendiculaires ; si le nombre entier *a* n'est pas supérieur au nombre entier *b*, c'est qu'il lui est inférieur ou égal ; etc. La logique qui sous-tend ainsi toute l'architecture mathématique traditionnelle est donc une logique binaire ; on dit encore qu'elle est bivalente ou qu'elle a deux valeurs de vérité, et deux seulement. Cela signifie simplement qu'une proposition est soit vraie, soit fausse et que, si elle n'est pas vraie, on peut être garant à coup sûr qu'elle est fausse (et vice versa).

Avec Riemann et Lobatchevski ◆, l'un posant que par un point pris hors d'une droite passent une infinité de parallèles et l'autre affirmant qu'il n'en passe aucune, le vrai est contraint de se relativiser. A savoir que ce qui est vrai dans une théorie pourra ne pas l'être dans une autre : on hésite à parler de vrai en soi. Et encore moins à justifier le vrai par le réel. Les mathématiques s'engagent dans la voie de l'axiomatisation. C'est-à-dire qu'est décrété vrai ce qui découle logiquement et non contradictoirement des définitions et des axiomes qu'on a posés. Mais, fondamentalement, le caractère bivalent de la logique n'est pas remis en cause. Il est même développé et formalisé par Boole ◆, notamment, à tel point que les logiciens sont encore victimes, au début du XIX[e] siècle, de la logique binaire considérée peu ou prou comme la reproduction non du monde, mais du mode de fonctionnement de l'esprit humain.

◆ B. Riemann (1826-1866) : mathématicien allemand qui ouvrit la voie à une nouvelle géométrie, non euclidienne, qui préparait le terrain pour la physique géométrique actuelle ; ses principaux travaux portent sur les fonctions abéliennes, la série hypergéométrique et les équations différentielles.
N. Lobatchevski (1793-1856) : mathématicien russe, surtout célèbre pour ses travaux sur les principes de la géométrie et sur la théorie des parallèles. Auteur de *Pangéométrie ou géométrie imaginaire* (1855), où il échafaude un nouveau système géométrique cohérent.

C'est dans le sillage du grand bouleversement mathématique contemporain que vont naître les logiques plurivalentes. Le point de départ en est simple : il consiste à ne plus retenir pour une proposition la seule alternative vrai/faux, mais à y adjoindre d'autres valeurs de vérité intermédiaires. C'est ce que fait en tout cas Lukasiewicz qui publie en 1920 les résultats et les principes de sa logique trivalente. Elle revient à admettre qu'entre le 1 de « c'est vrai » et le 0 de « c'est faux » il y a place pour le « c'est possible », auquel on pourra, par exemple, attribuer la valeur 1/2.

◆ Voir Algèbre de Boole.

Les mathématiciens en viennent aujourd'hui, à leur tour, à s'intéresser de très près à une façon plus souple, moins contraignante, d'appréhender leur objet. Il faut citer ici tout particulièrement les recherches entreprises, depuis 1965, par l'Américain L.A. Zadeh, professeur à l'université de Californie à Berkeley ◆.

◆ Recherches auxquelles un ouvrage tout récent de A. Kaufmann nous permet enfin d'avoir accès : *Introduction à la théorie des sous-ensembles flous.* Tome I : *Éléments théoriques de base* (Paris, Masson, 1973).

La théorie des ensembles flous

Elle s'inscrit directement dans le cadre de la théorie des ensembles classiques, où le degré d'appartenance d'un élément à un sous-ensemble est zéro — si l'élément n'appartient pas au sous-ensemble — ou un — s'il y appartient. Pas d'autre issue. Supposons maintenant que ce puisse être l'ensemble de tous les nombres compris entre 0 et 1, c'est-à-dire : 0 ou 1, mais aussi 0,2 ou 0,34 ou 0,957, etc. Bref, tous les nombres (qu'on dit « réels ») allant de 0 à 1. On notera cet ensemble : M = [0, 1] (à ne pas confondre avec $\{ 0, 1 \}$ qui, lui, est l'ensemble des deux nombres 0 et 1).

Tel est le point de départ sur lequel va s'édifier la notion de sous-ensemble flou. Les portes sont grand ouvertes. On trouve des degrés d'appartenance de 0,3, 0,7, 0,6, ce qui n'aurait aucun sens pour un mathématicien « binaire ».

L'intérêt, indique M. Kaufmann, est de permettre « la construction d'une structure » mathématique avec laquelle on pourra manipuler des concepts assez mal définis, » mais dont l'appartenance à des sous-ensembles a pu être hiérarchisée ». Et il en indique aussitôt un exemple : dans l'ensemble des hommes, soit le sous-ensemble flou des hommes de grande taille. L'appartenance à ce sous-ensemble sera « hiérarchisée » si, notamment, sa fonction d'appartenance est connue. Cette fonction, nous sommes entièrement libres de la déterminer nous-mêmes, compte tenu de l'exigence du problème que nous avons à traiter, c'est-à-dire ici de la façon dont nous entendons appréhender le phénomène « grande taille » dans une population donnée.

Par exemple, nous pouvons supposer que le maximum pour une « grande taille » est 2,5 mètres, et choisir alors la fonction d'appartenance $\mu(n) = \frac{x}{2,5}$ (n exprimant la taille en mètres). Cette formule est sans doute trop « large » et peu pertinente puisque, dès qu'un individu a une taille non nulle, même de 15 centimètres, alors il appartient d'une certaine façon au sous-ensemble flou des hommes de grande taille. Mais elle nous permet d'indiquer comment la hiérarchisation de l'appartenance peut s'effectuer. En effet, si André mesure 2 mètres, Jean 1 mètre et Pierre 1,8 mètre, leurs degrés d'appartenance sont respectivement de 0,8 ; 0,4 et 0,72. Si $\underset{\sim}{G}$ désigne le sous-ensemble flou (le tilde $\sim$ indique le sous-ensemble flou) des hommes de grande taille, nous écrivons donc :

$$\text{André} \underset{0,8}{\in} \underset{\sim}{G} \qquad \text{Jean} \underset{0,4}{\in} \underset{\sim}{G} \qquad \text{Pierre} \underset{0,72}{\in} \underset{\sim}{G}.$$

L'idée de départ consiste donc à relativiser, à coefficienter l'appartenance à un sous-ensemble. A définir une appartenance floue, donnant une signification à une expression telle que $x \underset{0,4}{\in} \underset{\sim}{G}$ (appartient au sous-ensemble flou $\underset{\sim}{G}$ avec le degré 0,4). Sur cette base, la théorie des sous-ensembles flous se développe en reprenant point par point les linéaments de la théorie des ensembles classique et en les transposant, une fois posées un certain nombre de définitions qui lui sont propres. C'est ainsi que sont introduites les notions d'inclusion floue, d'intersection floue, etc. : construction qui est aussi plus riche que la théorie des ensembles qu'elle prolonge, dans la mesure où elle paraît couvrir un champ plus vaste. Mais cet enrichissement, susceptible à coup sûr de multiples applications pratiques, correspond en fait, mathématiquement parlant, à un appauvrissement.

En effet, toutes les propriétés des sous-ensembles (ou parties) vulgaires ◆ se retrouvent dans les parties floues, sauf deux d'entre elles : l'intersection avec le complémentaire est égale à l'ensemble vide et la réunion est égale au référentiel E. Ce que l'on écrit $A \cap \overline{A} = \varnothing$ et $A \cup \overline{A} = E$. Ces deux propriétés ne sont plus valables pour une partie floue $\underset{\sim}{A}$.

◆ C'est le cas où l'ensemble des valeurs d'appartenance n'est plus [0, 1] mais se réduit au tandem (0, 1).

De ce fait, les propriétés du flou ne définissent plus une algèbre. Le calcul qui était possible avec les parties vulgaires ne l'est plus. La structure mathématique n'en existe pas moins. Mais elle est sensiblement plus pauvre que la structure « vulgaire » correspondante.

Les relations floues

Non seulement les relations de type classique sont généralisées (relations de préordre et d'ordre, notamment), mais d'autres plus complexes, plus floues sont définies. Ainsi, la relation de ressemblance ou celle de dissemblance. Elles ont

l'immense avantage de permettre une meilleure description de « beaucoup de » phénomènes complexes, réduits jusqu'à présent à des spécifications en tout » ou rien ».

Il n'y a pas de mathématique sans une logique qui la sous-tend. Et, en l'occurrence, à mathématique du flou, logique du flou.

La logique classique, c'est essentiellement George Boole qui lui a donné l'aspect formel que nous utilisons encore aujourd'hui. Cette logique est celle qui attribue deux valeurs de vérité exclusivement aux propositions et qui, à partir de là, permet par le jeu des « tables de vérité » de savoir quelle valeur il convient d'affecter à une proposition complexe.

Cette logique booléenne est directement associée à la théorie classique des ensembles. Il y a, en effet, équivalence terme à terme entre les opérations de la logique portant sur les propositions et les opérations ensemblistes. Il est aisé de vérifier, notamment, que le « et » de la logique correspond à l'intersection « ∩ » pour les ensembles.

De la même manière, la logique floue est associée à la théorie des sous-ensembles flous. C'est-à-dire qu'une fois définies l'intersection floue, la réunion floue et d'autres opérations élémentaires, on peut y rattacher la définition d'un « et » flou, d'un « ou » flou, etc. Par exemple, si les propositions floues p, q ont respectivement 0,3 et 0,7 comme valeur de vérité, on établit que la conjonction « $p \wedge q$ » (« p et q ») a 0,3 comme valeur (la plus petite des deux).

Si la logique booléenne permet un véritable calcul sur les propositions, en quoi elle est bien une algèbre, la logique floue ne le permettra pas. En particulier, il ne sera pas possible de dresser systématiquement des tables de vérité, comme on le faisait pour des variables binaires.

La logique floue et le réel

En ce qui concerne la logique floue, la « traduction » concrète sera inévitablement plus délicate (et plus coûteuse) puisqu'on ne pourra plus se contenter d'un matériel à deux positions : il faudra pouvoir obtenir et maintenir toutes les valeurs possibles entre 0 et 1. Cette technologie, néanmoins, existe d'ores et déjà (diode, triode, transistor, bascule, thyristor, etc.). Elle demande à être simplifiée, améliorée et rendue plus économique.

Notre pensée n'a pas accès au réel. Elle peut seulement en rendre compte par des modèles qui n'en sont que des approximations. Tout le processus de la connaissance s'organise donc autour du triplet suivant :

Pensée humaine ..... Modèle ..... Réel,

qui traduit approximativement l'inaccessibilité du réel en tant que tel pour la pensée.

Nul doute que, comme l'écrit le professeur Zadeh lui-même, la théorie des sous-ensembles flous soit un « pas vers un rapprochement entre la précision des mathé- » matiques classiques et la subtile imprécision du monde réel ».

Car la communication avec les ordinateurs, même les plus gros, reste très sommaire. Cette communication nous impose en effet une sorte de « binarisation » de notre pensée qu'assurent les langages symboliques courants (Fortran, Cobol, etc.).

Les tores de ferrite qui constituent la mémoire centrale ne sont susceptibles que de deux positions : soit ils sont « excités » (état 1), soit ils ne le sont pas (état 0). Or la logique et la mathématique floues autorisent un début de réponse à cette délicate question de la technologie d'aujourd'hui ◆.

◆ A. Kaufmann estime : « Lorsque seront opérationnels des softwares construits à partir d'une logique floue et, quand ce sera industriellement possible, des hardwares flous, alors la communication homme/machine sera beaucoup plus commode, rapide et mieux adaptée aux problèmes à résoudre. » « Software » désigne l'ensemble des éléments « intellectuels » de la machine : les programmes, le système qui lui permettent de « penser », par opposition à « hardware », qui est la « quincaillerie », la partie purement appareillage du calculateur.

Est-ce que la logique floue décrit exactement, formellement le fonctionnement de notre pensée?
C'est la tentation — idéaliste — à laquelle les mathématiciens du flou pourraient être tentés de succomber. Zadeh écrit en effet : « En vérité, on peut affirmer que » presque toute la logique du raisonnement humain n'est pas la logique classique » à deux valeurs ou même à plusieurs valeurs, mais une logique aux vérités floues, » aux conjonctions floues, aux règles de déduction floues. »
Et probablement, pour confirmer encore plus vigoureusement cette idée, va-t-on se mettre à rechercher les interrelations possibles entre la logique floue et le fonctionnement du système nerveux central. G. G.

FONCTION◆

pages 60, 61, 70, 100 à 152, 211, 231, 273, 276, 286, 291, 292, 301, 303, 304, 327 à 330, 341, 400, 413, 537.

◆ Lire aussi ce qui concerne la dérivée dans l'article « l'Analyse », page 125.

Pour étudier rapidement et efficacement une fonction, il est conseillé de procéder ainsi :

1) Préciser l'ensemble de définition de la fonction et ses intervalles de continuité. En général, elle est continue, ou, dans le cas le plus défavorable, elle a un nombre fini de points de discontinuité, et un nombre fini d'intervalles de continuité.

2) Étudier, dans ces intervalles, le signe de la dérivée et en conclure la variation de la fonction ; sur un intervalle de continuité :

— une dérivée positive entraîne une fonction croissante ;

— une dérivée négative entraîne une fonction décroissante ;

— une dérivée nulle en un point correspond à un extrémum de la fonction ;

— une dérivée nulle sur un intervalle entraîne une fonction constante sur cet intervalle.

3) Il ne reste plus qu'à évaluer les valeurs de la fonction en des points particuliers :

— valeur limite de la fonction aux extrémités des intervalles de continuité ;

— valeur de la fonction aux extrémums.

Il n'est pas interdit de calculer d'autres valeurs particulières de la fonction, ainsi que la dérivée en ces points.

4) Cette étude est généralement consignée dans un tableau à double entrée, appelé tableau de variation◆, permettant une vision rapide de l'allure de la fonction : sur une ligne horizontale, on indique le domaine de définition de la fonction, en n'oubliant pas de noter tous les points de discontinuité. Sous ces points de discontinuité, on trace deux lignes verticales ; sur une ligne verticale, on indique d'abord la dérivée et, sur la ligne horizontale correspondante, on note le signe ou la nullité de la dérivée. Pour la fonction, on trace alors des traits montant ou descendant suivant que la dérivée est positive ou négative, à l'intérieur de chaque intervalle de continuité. On indique également les valeurs particulières, calculées pour cette fonction. On termine en général par le tracé de la courbe représentative de cette fonction.

◆ Voir page suivante.

Exemple : étude de la fonction $f(x) = x^2$ (fonction puissance 2)

1) L'ensemble de définition est ici l'ensemble des nombres réels : $\mathbb{R}$. La fonction est continue sur tout cet ensemble.

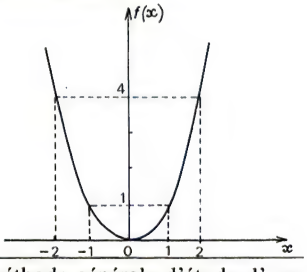
2) Dérivée : $f'(x) = 2x$◆.
Cette dérivée est positive si x est positif, nulle si x est nul, négative si x est négatif.

◆ Voir page 125.

3) Limites :
— quand x tend vers $+\infty$, x^2 tend vers $+\infty$;
— quand x tend vers $-\infty$, x^2 tend vers $+\infty$.
Valeurs particulières :
— quand $x = 0$, $f(x) = 0$;
— quand $x = 1$ ou $x = -1$, $f(x) = 1$;
— quand $x = 2$ ou $x = -2$, $f(x) = 4$.
4) Tableau de variation :

x	$-\infty$		0		$+\infty$
$f'(2)$		$-$	0	$+$	
$f(x)$	$+\infty$	↘	0	↗	$+\infty$

Courbe représentative

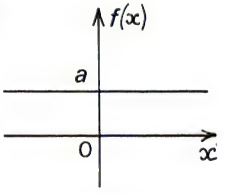


Il est évident que nous venons d'indiquer ici la méthode générale d'étude d'une fonction. Pour certaines fonctions particulières, toute autre méthode plus directe n'est pas à exclure ; en voici quelques exemples.

La fonction constante

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \longmapsto f(x) = a,$$

où a est un nombre réel fixe, a la représentation graphique ci-contre.

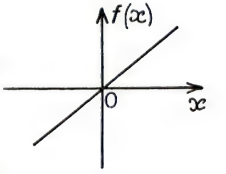


$f(x) = d$

La fonction identité

$$I : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \longmapsto I(x) = x$$

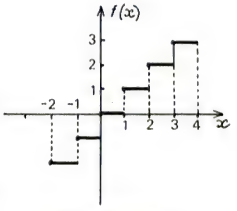
a un graphe coïncidant avec la bissectrice de l'angle Ox, Oy, puisque tout point M du graphe a pour coordonnées (x, x).



$f(x) = x$

La fonction partie entière de x

De même, l'étude de la fonction partie entière de x se fait directement. A tout nombre réel x, elle associe sa partie entière (la partie entière de 3,25 est 3 ; de même, la partie entière de tout nombre réel supérieur ou égal à 3 et inférieur à 4 est le nombre entier 3), d'où le graphe représenté ci-contre.



partie entière de x

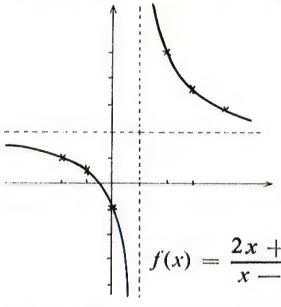
La fonction homographique

Fonction réelle f d'une variable réelle x, qui à x fait correspondre $f(x)$ de la forme $\dfrac{ax + b}{cx + d}$ (où a, b, c, d sont des nombres réels non tous nuls).

● Si les quatre nombres a, b, c, d sont tels que $ad - bc$ soit différent de zéro, la fonction homographique est dite *propre*. Sa courbe représentative est une hyperbole (dont les asymptotes sont les droites d'équations $y = \dfrac{a}{c}$ et $x = -\dfrac{d}{c}$) ◆.

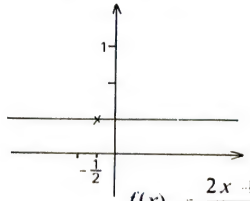
◆ Voir graphiques page suivante.

Fonction homographique propre



$$f(x) = \frac{2x + 1}{x - 1}$$

Fonction homographique dégénérée



$$f(x) = \frac{2x + 1}{6x + 3} = \frac{2x + 1}{3(2x + 1)} = \frac{1}{3} \text{ si } 2x + 1 \neq 0$$

● Si a, b, c, d sont tels que $ad - bc = 0$, la fonction est dite *dégénérée*. Sa courbe représentative est une droite parallèle à Ox, privée du point d'abscisse $x = -\frac{d}{c}$ (car, pour cette valeur, le dénominateur de la fraction est nul).

La fonction cumulative : voir l'article « les Probabilités », page 428.

La fonction logarithme : voir Exponentielle (fonction) et Logarithme.

La fonction de répartition : voir l'article « les Probabilités », page 427.

Les fonctions trigonométriques : voir Trigonométrie.

Tableau de dérivées

Fonction	Dérivée
$f(x) = a$	$f'(x) = 0$
$f(x) = x$	$f'(x) = 1$
$f(x) = x^2$	$f'(x) = 2x$
$f(x) = \alpha x^2 + \beta x + \gamma$	$f'(x) = 2\alpha x + \beta$
$f(x) = x^3$	$f'(x) = 3x^2$
$f(x) = x^{\frac{1}{2}}$	$f'(x) = \frac{1}{2x^{\frac{1}{2}}}$
$f(x) = x^{\frac{1}{3}}$	$f'(x) = \frac{1}{3x^{\frac{2}{3}}}$
$\alpha \neq 0 \quad f(x) = x^\alpha$	$f'(x) = \alpha x^{\alpha-1}$
$f(x) = \cos x$	$f'(x) = -\sin x$
$f(x) = \sin x$	$f'(x) = \cos x$
$f(x) = \operatorname{tg} x$	$f'(x) = \frac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x$
$f(x) = \operatorname{Log} x$	$f'(x) = \frac{1}{x}$
$f(x) = e^x$	$f'(x) = e^x$

Tableau de primitives

Fonction	Primitive
$f(x) = a$	$F(x) = ax + C$
$f(x) = x$	$F(x) = \frac{x^2}{2} + C$
$f(x) = x^2$	$F(x) = \frac{x^3}{3} + C$
$f(x) = \alpha x^2 + \beta x + \gamma$	$F(x) = \frac{\alpha x^3}{3} + \frac{\beta x^2}{2} + \gamma x + C$
$f(x) = x^3$	$F(x) = \frac{x^4}{4} + C$
$f(x) = x^{\frac{1}{2}}$	$F(x) = \frac{2}{3}x^{\frac{3}{2}} + C$
$f(x) = x^{\frac{1}{3}}$	$F(x) = \frac{3}{4}x^{\frac{4}{3}} + C$
$\alpha \neq -1 \quad f(x) = x^\alpha$	$F(x) = \frac{x^{\alpha+1}}{\alpha + 1} + C$
$f(x) = \cos x$	$F(x) = \sin x$
$f(x) = \sin x$	$F(x) = -\cos x$
$f(x) = \operatorname{tg} x$	$F(x) = -\operatorname{Log} \lvert\cos x\rvert$
$f(x) = \frac{1}{x}$	$F(x) = \operatorname{Log} x$
$f(x) = e^x$	$F(x) = e^x$

C = constante réelle arbitraire

FONCTIONNELLE (Relation) ◆

◆ Lire d'abord Relation.

Prenons un exemple : la relation « a pour carré » définie de l'ensemble $E = \{ -1, 1, 2, 3 \}$ vers l'ensemble $F = \{ 1, 2, 3, 4 \}$ est telle que :

— 1 a pour carré 1
1 a pour carré 1
2 a pour carré 4
3 n'a pas de carré dans F.

Relation « a pour carré ».

Relation fonctionnelle.

Chaque élément de E est en relation avec au plus un élément de F ; c'est une relation fonctionnelle. Par contre, la relation « est le carré de » définie de $E = \{ 1, 2, 3, 4 \}$ vers $F = \{ -2, -1, 0, 1, 2 \}$ n'est pas fonctionnelle, car 1, élément de E, est le carré de 1, mais aussi de — 1 ; de même, 4 est le carré de 2, mais aussi de — 2. Les deux éléments de E 1 et 4 sont chacun en relation avec deux éléments de F.

Relation « est le carré de ».

Relation non fonctionnelle

Une relation $\mathcal{R}$ d'un ensemble E vers un ensemble F est dite fonctionnelle si chaque élément x de E a au plus une image par $\mathcal{R}$ dans F.

Le diagramme sagittal d'une relation fonctionnelle est tel que, de tout élément de E, part une flèche au plus, aboutissant à un élément de F (de certains éléments de E, il part une flèche ; d'autres éléments, il n'en part pas).
Si la relation $\mathcal{R}$ est une relation fonctionnelle de E vers F, on abandonne la notation $x \mathcal{R} y$ (indiquant que x est en relation par $\mathcal{R}$ avec y) pour la notation $y = f(x)$ (on lit : égale f de x), dans laquelle f représente la relation fonctionnelle. On emploie également la notation :

$$f : x \longmapsto y \text{ (ou } f : x \rightarrowtail y)$$

(on lit : f fait correspondre, à x, y ou : x a pour image, par f, y).

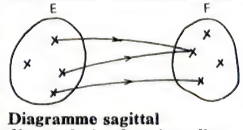


Diagramme sagittal d'une relation fonctionnelle.

FORME ◆ pages 78 et 79.

◆ Lire d'abord l'article « l'Algèbre linéaire » pour la définition d'espace vectoriel.

Si E est un espace vectoriel sur un corps K (qui peut être en particulier le corps $\mathbb{R}$ des nombres réels ou le corps $\mathbb{C}$ des nombres complexes), on appelle forme sur E toute application de E dans K, c'est-à-dire toute application qui associe à un élément de E un élément de K.
Par exemple, si E est l'ensemble $\mathbb{R}^2$ des couples (x, y) de nombres réels ◆, l'application qui au couple (x, y) fait correspondre le nombre réel $x + y$ est une forme sur $\mathbb{R}^2$.

◆ $\mathbb{R}^2$ est un espace vectoriel sur $\mathbb{R}$.

Forme linéaire

Une forme f sur un espace vectoriel E est dite linéaire si, quels que soient les vecteurs $\vec{u}$ et $\vec{v}$ de E et quel que soit le scalaire λ de K, on a :

$$\begin{cases} f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}) \\ f(\lambda . \vec{u}) = \lambda . f(\vec{u}) \end{cases}$$

La forme ci-dessus, définie par $f(\vec{u}) = x + y$ si $\vec{u} = (x, y)$, est une forme linéaire, car :

$$\left.\begin{array}{l}\text{si } \vec{u} = (x, y)\\ \vec{v} = (x', y')\end{array}\right\} \vec{u} + \vec{v} = (x + x', y + y')$$

$$\left.\begin{array}{l} f(\vec{u}) = x + y\\ f(\vec{v}) = x' + y'\end{array}\right\} f(\vec{u}) + f(\vec{v}) = \underbrace{(x + y) + (x' + y')}_{(1)}$$

$$f(\vec{u} + \vec{v}) = \underbrace{(x + x') + (y + y')}_{(2)}$$

(1) et (2) sont égaux grâce à la commutativité et à l'associativité de l'addition dans $\mathbb{R}$.

si $\vec{u} = (x, y)$ alors $\lambda.\vec{u} = (\lambda.x, \lambda.y)$

$$\left.\begin{array}{l} f(\lambda.\vec{u}) = \lambda.x + \lambda.y\\ \lambda.f(\vec{u}) = \lambda.(x + y)\end{array}\right\} \text{Ces deux nombres sont égaux grâce à la distributivité de la multiplication sur l'addition dans } \mathbb{R}.$$

Forme bilinéaire

C'est une application qui fait correspondre au couple de vecteurs $\vec{u}$ et $\vec{v}$ de E un élément de K (c'est une application de E × E dans K) et qui vérifie les deux propriétés suivantes :

- Elle est linéaire par rapport à $\vec{u}$:

$$f(\vec{u_1} + \vec{u_2}, \vec{v}) = f(u_1, \vec{v}) + f(\vec{u_2}, \vec{v})$$
$$f(\lambda.\vec{u}, \vec{v}) = \lambda.f(\vec{u}, \vec{v}).$$

- Elle est linéaire par rapport à $\vec{v}$:

$$f(\vec{u}, \vec{v_1} + \vec{v_2}) = f(\vec{u}, \vec{v_1}) + f(\vec{u}, \vec{v_2})$$
$$f(\vec{u}, \lambda.\vec{v}) = \lambda.f(\vec{u}, \vec{v}).$$

Par exemple, l'application qui au couple $(\vec{u}, \vec{v})$ de vecteurs de $\mathbb{R}^2$:

$$\vec{u} = (x, y) \quad \text{et} \quad \vec{v} = (x', y')$$

fait correspondre le nombre réel $xx' + yy'$ est une forme bilinéaire.
Une forme bilinéaire est dite :

— *symétrique* si, quels que soient $\vec{u}$ et $\vec{v}$, $f(\vec{u}, \vec{v}) = f(\vec{v}, \vec{u})$;

— *antisymétrique* si, quels que soient $\vec{u}$ et $\vec{v}$, $f(\vec{u}, \vec{v}) = -f(\vec{v}, \vec{u})$.

Forme multilinéaire

Une forme multilinéaire est une application qui, à n vecteurs $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}$ de E, fait correspondre un élément de K et qui est linéaire par rapport à chacun des vecteurs $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}$.
Une telle forme est dite *alternée* si sa valeur est changée de signe lorsqu'on échange deux vecteurs quelconques :

$$f(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_i}, \ldots, \vec{u_j}, \ldots, \vec{u_n}) = -f(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_j}, \ldots, \vec{u_i}, \ldots, \vec{u_n}).$$

Forme quadratique

Étant donné une forme bilinéaire symétrique f sur l'espace vectoriel E, on appelle forme quadratique associée l'application Q (de E dans K) qui à tout vecteur $\vec{u}$ de E fait correspondre l'élément de K noté $Q(\vec{u})$, défini par :

$$Q(\vec{u}) = f(\vec{u}, \vec{u}).$$

Par exemple, si on a défini sur $\mathbb{R}^2$ la forme bilinéaire symétrique f :

$$f(\vec{u}, \vec{v}) = xx' + yy',$$

la forme quadratique associée Q est définie par :

$$Q(\vec{u}) = f(\vec{u}, \vec{u}) = x^2 + y^2.$$

Si K est le corps $\mathbb{R}$ des nombres réels, une forme quadratique est définie positive si, quel que soit le vecteur $\vec{u}$ de E, le nombre réel $Q(\vec{u})$ est positif (strictement si $\vec{u}$ n'est pas nul).

Par exemple, la forme quadratique Q, définie sur $\mathbb{R}^2$, par :

$$Q(\vec{u}) = x^2 + y^2,$$

est définie positive, car $x^2 + y^2$ ne peut être nul (dans $\mathbb{R}$) que si $x = y = 0$, c'est-à-dire si $\vec{u} = \vec{0}$; pour tous les autres vecteurs, $x^2 + y^2$ est positif strictement.

FOYER : voir Coniques

FRACTIONNAIRE (Nombre) : voir Rationnel (Nombre) et les pages 238, 368 et 369.

FRÉQUENCE DE CLASSE : voir l'article « les Statistiques », pages 460, 485 et la page 176.

FRONTIÈRE ◆

◆ Vocabulaire de topologie. Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

Un intervalle quelconque, non vide, (a, b) de la droite numérique $\mathbb{R}$ (ensemble des nombres réels) admet :

— *pour intérieur :* l'intervalle ouvert $]a, b[$;

— *pour adhérence :* l'intervalle fermé $[a, b]$;

— *pour frontière :* l'ensemble constitué par les deux points a et b : $\{a, b\}$.

La frontière d'un sous-ensemble A d'un espace topologique E est l'ensemble $\overline{A} - \mathring{A}$ (adhérence de A moins intérieur de A). Elle est aussi égale à l'intersection de l'adhérence de A avec l'adhérence du complémentaire de A dans E : $\overline{A} \cap \overline{\complement_E A}$.

GH

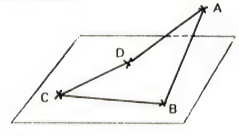
GALOIS (Corps de)◆

◆ Lire d'abord, dans l'article « l'Algèbre », le paragraphe « Corps ». Congruence.

On appelle corps de Galois un corps ayant un nombre fini d'éléments.
Par exemple, $\mathbb{Z}/p$, pour p entier naturel premier, est un corps qui a p éléments ; c'est un corps de Galois.

GAUCHE (Courbe, Quadrilatère) pages 328, 525.

Quadrilatère gauche



Une courbe gauche est une courbe de l'espace, non plane.
Une hélice, par exemple, est une courbe gauche.
Un quadrilatère gauche est un quadrilatère dont les quatre sommets ne sont pas situés dans un même plan.

GAUSS (Courbe de, Loi de) : voir l'article « les Statistiques », pages 468 à 471, 494 à 505.

GÉNÉRATEUR : voir l'article « l'Algèbre linéaire », pages 65 à 67, 79 et les pages 33 et 34.

GÉNÉRATRICE : voir Cône, Cylindre, Surface de révolution pages 225, 520.

GÉOMÉTRIE pages 55, 63, 186, 191 à 194, 203, 231, 324, 326 à 330, 526, 528.

A l'origine, la géométrie est une science expérimentale ; née de la nécessité de mesurer des longueurs, des aires, des volumes, etc., son objet d'étude est l'espace physique qui nous entoure (espace à deux dimensions dans le cas de la géométrie plane, à trois dimensions pour la géométrie dans l'espace).
La géométrie étudie les points et les figures qu'ils forment : droites, plans, etc. A la base de cette étude, il y a un système d'axiomes, à partir desquels des raisonnements logiques permettent de déduire d'autres propriétés. La première tentative de construction rigoureuse de la géométrie est due à Euclide ; parmi les axiomes de départ figure le célèbre postulat d'Euclide : « Par un point situé hors d'une droite, il passe une parallèle à cette droite et une seule. » Au XIX^{e} siècle, Gauss et Riemann construisirent des théories cohérentes où le postulat d'Euclide était remplacé par un autre axiome : pour Gauss, Lobatchevski et Bolyai, « par un point extérieur à une droite, il passe une infinité de parallèles à cette droite » (cette géométrie est appelée géométrie hyperbolique) ; pour Riemann, « par un point extérieur à une droite, il ne passe aucune parallèle à cette droite » (cette géométrie est appelée géométrie elliptique). Ces géométries sont qualifiées de géométries non euclidiennes. Ce ne sont pas simples jeux de l'esprit et leurs applications sont nombreuses◆.
Au cours de la seconde moitié du XIX^{e} siècle, la notion de groupe est formulée de manière précise et l'on s'aperçoit que les théorèmes de géométrie ne sont que

◆ Notamment en relativité générale pour les espaces de Riemann.

les formulations en un autre langage des théorèmes d'une théorie plus générale : l'algèbre linéaire.

> Actuellement, le sens donné au mot « géométrie » est le suivant : une géométrie consiste en l'étude d'un ensemble et des propriétés invariantes par certaines transformations portant sur cet ensemble.

La géométrie projective plane est l'étude des propriétés conservées par une transformation projective, c'est-à-dire une transformation qui, à tout point M du plan de coordonnées homogènes (x, y, t)◆, fait correspondre le point M′ de coordonnées (x', y', t') données par les formules :

◆ Voir Coordonnées.

$$\begin{cases} x' = ax + by + ct + x_0 \\ y' = a'x + b'y + c't + y_0 \\ t' = a''x + b''y + c''t + t_0 \end{cases}$$

où la matrice $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$ est une matrice régulière.

De nombreuses propriétés d'une figure géométrique sont conservées par une transformation projective ; ces propriétés sont dites projectives. En particulier, le birapport de quatre points ou de quatre droites est conservé

La géométrie affine plane est l'étude des propriétés conservées par une transformation affine, c'est-à-dire une transformation qui, à tout point M du plan, de coordonnées homogènes (x, y, t), fait correspondre le point M′ de coordonnées (x', y', t') données par les formules :

$$\begin{cases} x' = ax + by + ct \\ y' = a'x + b'y + c't \\ t' = t \end{cases}$$

où la matrice $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ 0 & 0 & 1 \end{pmatrix}$ est telle que $ab' - ba'$ ne soit pas nul. Les propriétés conservées par une telle transformation sont dites propriétés affines. Le birapport est conservé, mais aussi les notions de milieu d'un segment, de parallélisme de deux droites.

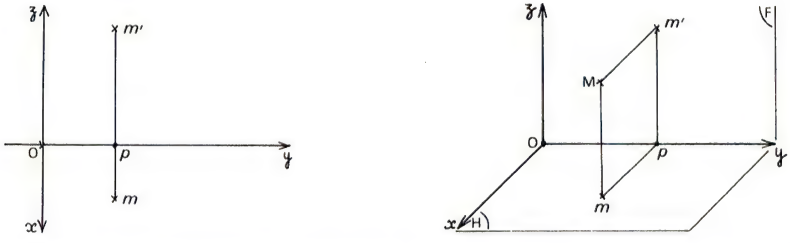
La géométrie métrique est l'étude des propriétés conservées par une transformation métrique, c'est-à-dire une transformation affine particulière, pour laquelle la matrice $\begin{pmatrix} a & b \\ a' & b' \end{pmatrix}$ est une matrice orthogonale.

Les propriétés invariantes sont appelées propriétés métriques : le produit scalaire de deux vecteurs est conservé, les longueurs, les angles également.

Géométrie descriptive

La géométrie descriptive est l'étude de figures simples de l'espace à l'aide d'une représentation plane particulière de cet espace (cette représentation est due à Monge).

Représentation plane **Figure dans l'espace**



Un point M de l'espace est repéré par ses deux projections m et m' sur deux plans perpendiculaires, un plan horizontal H ◆ et un plan vertical F ◆. La droite d'intersection de ces deux plans est appelée ligne de terre.

◆ Appelé plan horizontal de référence.

◆ Appelé plan frontal de référence.

Géométrie analytique

Étude des ensembles de points de la géométrie élémentaire à l'aide du calcul algébrique. Chaque point est représenté par un ensemble de nombres réels : ses coordonnées. Un problème de géométrie analytique se traite en trois étapes :

— *1re étape :* traduire en langage algébrique les données géométriques du problème ; c'est ce qu'on appelle la mise en équations du problème.

— *2e étape :* résoudre le problème algébrique ainsi formulé.

— *3e étape :* traduire en langage géométrique les résultats fournis par le calcul algébrique.

Géométrie différentielle ou infinitésimale

Étude précise des courbes et des surfaces grâce au calcul algébrique. La géométrie infinitésimale fait très largement appel aux notions de dérivées, de dérivées partielles, de différentielles, etc. Elle étudie notamment la courbure, les tangentes, les normales à une surface, etc.

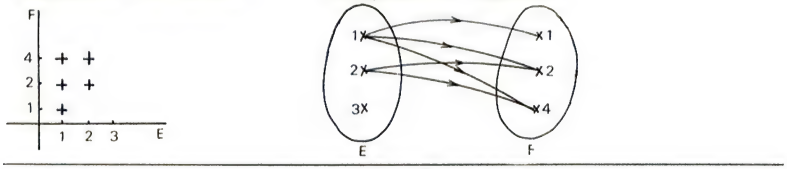
GRAPHE (d'une relation) pages 169, 174, 274.

Si $\mathcal{R}$ est une relation d'un ensemble E vers un ensemble F, on appelle graphe de $\mathcal{R}$ l'ensemble des couples (x, y) (x élément de E et y élément de F) qui vérifient la relation $\mathcal{R}$.

Exemple : $E = \{1, 2, 3\}$, $F = \{1, 2, 4\}$. La relation $\mathcal{R}$ étant la relation « x divise y », le graphe de $\mathcal{R}$ est : $\{(1, 1), (1, 2), (1, 4), (2, 2), (2,4)\}$

Le graphe d'une relation $\mathcal{R}$ peut être représenté au moyen d'un diagramme cartésien ou d'un diagramme sagittal.

Diagramme cartésien **Diagramme sagittal**

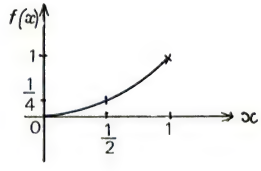


Dans le cas d'une application f d'un ensemble E vers un ensemble F, le graphe de l'application f est constitué par les couples (x, y) (x appartenant à E et y à F) tels que y soit l'image de x par f. On peut représenter également le graphe d'une application soit par un diagramme sagittal, soit par un diagramme cartésien (le diagramme ayant ceci de particulier que chaque élément x de E a une image et une seule dans F).

Exemple : Si E est l'intervalle [0, 1] ◆ et si l'application f est celle qui à tout nombre réel x fait correspondre son carré x^2 : $f : x \longmapsto x^2$.

Le graphe de cette application peut être représenté par la courbe ci-contre.

◆ Voir l'article « l'Analyse » et Intervalle.



GRAVITÉ (Centre de) : voir Médiane

GROSSIÈRE (Topologie) : voir Topologie

GROUPE COMMUTATIF, CYCLIQUE, SYMÉTRIQUE :

voir l'article « l'Algèbre », pages 28 à 45, 33 à 35, et les pages 57, 85, 102, 166, 216, 329, 368, 369, 370, 373, 394.

HARMONIQUE (Fonction) : voir Équations aux dérivées partielles

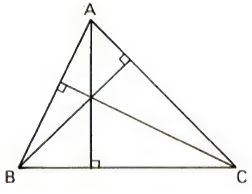
HAUTEUR

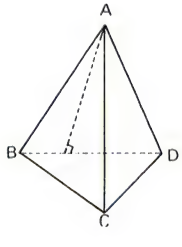
Hauteur d'un triangle

Si ABC est un triangle, la hauteur issue de A est la droite passant par A et perpendiculaire au côté BC. On définit de même les hauteurs issues de B et de C. Ces trois droites se coupent en un même point, appelé *orthocentre* du triangle.

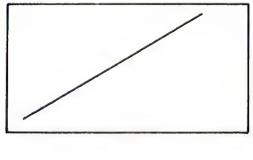
Hauteur d'un tétraèdre

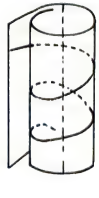
Chaque droite issue d'un sommet et perpendiculaire à la face opposée.





HÉLICE CIRCULAIRE



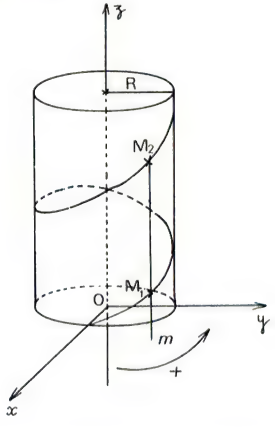


On peut obtenir aisément une hélice circulaire : il suffit de tracer sur un rectangle de papier une droite (non parallèle aux côtés du rectangle) puis d'enrouler ce rectangle sur lui-même de façon à former un cylindre.

Une hélice circulaire est une courbe tracée sur un cylindre de révolution, telle qu'en chacun de ses points sa tangente forme avec Oz un angle de mesure constante. Les équations paramétriques d'une telle courbe sont :

$$\begin{cases} x = R \cos t \\ y = R \sin t \\ z = ht \end{cases}$$

où R est le rayon du cylindre et h un nombre constant plus grand que zéro. Le point m, projection sur xOy du point M de l'hélice, décrit le cercle de centre O et de rayon R. Lorsque le paramètre t augmente de 2π, le point m fait un tour complet du cercle dans le sens positif. Le point M s'est déplacé sur l'hélice de la position M_1 à la position M_2.



La droite M_1M_2 est parallèle à Oz et la longueur de M_1M_2 est égale à $2\pi h$. Cette distance est appelée le *pas de l'hélice*.
Remarques : si $h > 0$, l'hélice s'enroule dans le sens direct ; il arrive qu'on parle d'hélice directe. Si $h < 0$, l'hélice s'enroule dans le sens rétrograde ; il arrive alors qu'on parle d'hélice rétrograde.

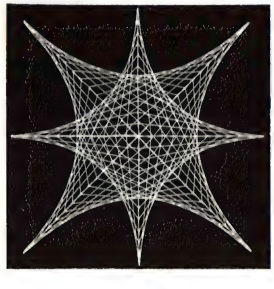
HENRY (Droite de) : voir l'article « les Statistiques », page 503.

HEPTAGONE : voir Polygones

HEXAGONE : voir Polygones

HISTOGRAMME : voir l'article « les Statistiques », pages 461, 469.

HISTORIQUE : voir l'article pages suivantes



Historique

par François Russo

L'Antiquité, en dehors du monde grec et romain

Il est malaisé de présenter un tableau satisfaisant de l'histoire des mathématiques dans l'Antiquité en raison de la grande diversité des peuples et des civilisations qui s'y rencontrent, souvent enchevêtrés, et dont l'histoire porte sur plus de deux millénaires, et aussi du fait de la rareté des documents. De toute façon, comme il s'agit d'une mathématique assez rudimentaire, nous pouvons limiter sa description à quelques traits essentiels. Le champ des mathématiques est à cette époque assez restreint, surtout constitué par les systèmes de figuration des nombres et les méthodes de calcul que prolonge un début d'algèbre, puis par une géométrie très sommaire. D'une manière générale, il s'agit de mathématiques avant tout pratiques (comptabilité, métrologie, mesure de volumes et de surfaces, calculs astronomiques). C'est seulement, semble-t-il, en Mésopotamie que les mathématiques offrent des aspects un peu désintéressés.

François Russo, s. j.
Né en 1909. Polytechnicien ; historien et philosophe des sciences.
Membre correspondant de l'Académie internationale d'histoire des sciences, il est également l'auteur de plusieurs ouvrages, dont : *Bibliographie d'histoire des sciences et des techniques* (Paris, Hermann, 1969), *Pour une bibliothèque scientifique* (Paris, Le Seuil, 1972).

La Grèce

Si notable que soit la contribution de l'Antiquité au progrès des mathématiques en dehors de la Grèce, l'apport de celle-ci est incomparablement supérieur. Pour comprendre ce « miracle », il convient de ne pas isoler l'histoire des mathématiques de l'histoire générale de la pensée. La création mathématique grecque procède essentiellement d'un idéal de pensée désintéressé, d'un souci de rationalité, de logique, dont Aristote devait apporter la formulation la plus élaborée. On comprend alors la rigueur et l'esprit de synthèse qui marquent les mathématiques grecques.

L'idéal mathématique grec s'est d'abord traduit dans la réflexion sur le nombre. Avec Pythagore (585-500 av. J.-C.), pour la première fois est reconnu le nombre irrationnel. Il s'offre naturellement lorsque l'on veut calculer la diagonale du carré. Les Grecs démontrent sans peine que cette diagonale ne peut s'exprimer en fraction rationnelle du côté. Mais la réflexion grecque sur le nombre ne se limitera pas aux quantités incommensurables dont le livre VII des *Éléments* d'Euclide (III^e siècle av. J.-C.) nous fournit une étude développée. Fruit d'une longue élaboration où Eudoxe a dû jouer le rôle principal, la théorie des grandeurs, au livre V des *Éléments*, est d'une qualité telle qu'il

faudra attendre le début du XIXe siècle pour retrouver une réflexion de même niveau.
Fondée sur la notion de rapport de deux grandeurs quelconques (et non pas de nombres entiers), elle est développée de telle façon qu'à quelques retouches près on peut y voir une présentation satisfaisante pour la pensée mathématique moderne. Nous noterons la place importante qu'a tenue dans les mathématiques grecques le nombre entier, en liaison d'ailleurs avec des vues philosophiques, notamment chez Pythagore et Platon. Très liée à la géométrie, la spéculation sur les nombres a fait une grande place aux nombres dits figurés, constitués par des points formant des figures simples (cercle, carré, triangle, cube).
La théorie proprement dite des nombres entiers occupe le livre VIII des *Éléments* d'Euclide. Mais les résultats sont tous élémentaires. Après Euclide, on trouve certains résultats nouveaux. Notons, à la fin du IIIe siècle av. J.-C., le « crible » d'Eratosthène (284-192 av. J.-C.) pour la détermination des nombres premiers et, beaucoup plus tard, une renaissance du pythagorisme avec Nicomaque de Gérase.

Le Moyen Age, du VIe au XIVe siècle

Cette période de huit siècles, qui s'étend du VIe au XIVe siècle, ne nous offre aucun progrès fondamental dans le développement des mathématiques. L'apport le plus notable est celui des Arabes qui, outre des contributions originales de valeur, surtout dans le domaine de l'algèbre, jouèrent le rôle de relais de la science grecque surtout, mais aussi de celle de l'Inde et, sans doute aussi, de celle de la Chine.
Du IXe au XIIe siècle, en dehors de Gerbert — le pape Sylvestre II (930-1003) —, les mathématiciens dont le nom mérite d'être retenu appartiennent à l'islam : au IXe siècle, Al Khwarismi Abu Kamil ; au Xe, Al Battani; au XIe, Al Kharki, Al Zarqali; au XIIe, Al Bitrugi. Il faut en outre mentionner pour l'Inde, au XIIe siècle, Bhaskara (1114-1183). Plus tard, au contraire, les personnages majeurs, mais de moindre envergure que les précédents, appartiennent au monde chrétien : au XIIIe siècle, Léonard de Pise, Sacrobosco, Jordanus de Nemore, Campanus de Novare, Robert Grosseteste ; au XIVe siècle, Bradwardine, Nicolas Oresme.

La Renaissance, du XVe au XVIe siècle

Pour l'histoire des mathématiques, comme d'ailleurs pour l'histoire générale, cette période se divise pour la plupart en écoles nationales ayant chacune son style propre : l'école allemande, plus formaliste que créatrice, avec Peurbach (1425-1461), Regiomontanus (1466-1476), Rudolff (début du XVIe siècle), Stifel (1486-1567), le jésuite Clavius (1538-1612) ; l'école italienne avec Luca Pacioli (1445-1517), Scipione del Ferro (1465-1526), Tartaglia (1505-1557), Cardan (1501-1576), Bombelli (1530-?) ; l'école anglaise avec Recorde (1510-1588), Neper (1550-1617), Hariott (1560-1621).
Pour les autres pays, on ne peut parler vraiment d'écoles. En France,

deux noms s'imposent : au début de la Renaissance, Nicolas Chuquet (1445-1500) et, à la fin, Viète (1540-1603). Il faut aussi mentionner le Portugais Nunez (1492-1577) et le Hollandais Stevin (1548-1620).

XVII^e siècle

Après les progrès estimables mais limités des mathématiques au Moyen Age et à la Renaissance, c'est un grand moment de création que le XVII^e siècle nous offre. Cinq grands noms dominent : dans la première moitié du siècle, Descartes (1596-1650), Fermat (1601-1665), Pascal (1623-1662). Dans la seconde, Leibniz (1646-1716) et Newton (1642-1727). Viennent ensuite dans un rang honorable quelque vingt noms ; par ordre à peu près chronologique : Briggs, Kepler, Cavalieri, Roberval, Mersenne, Desargues, Schooten, Sluse, Grégoire de Saint-Vincent, Torricelli, Hudde, Wallis, Barrow, Huygens, J. Gregory, Tchirnhaus, Jacques et Jean Bernoulli.
La grande affaire est alors la création du calcul infinitésimal. Mais en d'autres domaines se rencontrent aussi des progrès remarquables, principalement en géométrie projective, en géométrie analytique et en théorie des nombres.

XVIII^e siècle

En mathématiques, comme dans les autres domaines d'ailleurs, la période qui s'étend de 1715 environ jusqu'à la fin de la Révolution française n'est pas un moment de grande création. Nous y verrons une période de transition qui exploite, fort intelligemment d'ailleurs, les conquêtes du XVII^e siècle, surtout dans le domaine du calcul infinitésimal, qui s'ouvrent sans doute à d'importantes perspectives nouvelles, mais sans atteindre aux démarches décisives qui, au siècle suivant, leur donneront toute leur signification.
Un nom domine cette époque, celui de l'Allemand Euler (1707-1783), dont l'œuvre, qui s'étend à tous les domaines de la mathématique, culmine au milieu du siècle.
Deux autres noms doivent ensuite être cités : d'Alembert (1177-1783) et Lagrange (1736-1813). Le premier, dont l'œuvre se situe avant la Révolution dans la période de l'*Encyclopédie*, est à la fois algébriste et analyste. Le second, avant tout algébriste, appartient à la période de la Révolution et de l'Empire, et son style annonce plus directement le XIX^e siècle. On mentionnera ensuite, dans la première moitié du siècle, Daniel Bernoulli (1700-1782), Clairaut (1713-1765) et Maclaurin (1698-1746) ; et, dans la seconde, Monge (1746-1818), dont l'œuvre déborde sur le XIX^e siècle, sans doute surtout astronome et « mécanicien », mais dont l'apport mathématique est aussi remarquable.

XIX^e et XX^e siècle

C'est l'époque du grand essor des mathématiques. On en fixera environ le début à l'issue de la Révolution française. Alors va s'affir-

mer un style nouveau des mathématiques, marqué essentiellement par un souci de plus grande rigueur, par un esprit de conquête — qui conduira à l'ouverture de nombreux champs nouveaux d'investigation — et par la recherche d'une généralité toujours plus grande. Ainsi les mathématiques reposeront sur des bases plus solides et leur véritable caractère apparaîtra avec le développement de l'axiomatisation et l'introduction du concept de structure abstraite. Les mathématiques, jusque-là assez disparates, se réorganiseront alors selon des perspectives plus fondamentales qui conduiront sinon à son unification complète, du moins à une claire mise en évidence de la vraie nature de ses différents domaines et de leurs relations.

Dès le début du XIX^e siècle, la production mathématique connaît une grande expansion, comme le montre notamment la création de plusieurs périodiques mathématiques tels que le *Journal* de Gergonne (1810-1832), le *Journal de l'École polytechnique* (1794), le *Journal* de Crelle (1826), le *Journal* de Liouville (1836). Dans la seconde moitié du XIX^e siècle seront créées les *Annales de l'École normale supérieure* (1864), les *Mathematische Annalen* (1868).

Quant aux mathématiciens eux-mêmes, ils sont légion. « Nous aurons, » écrit Félix Klein, un tableau du développement des mathématiques » si nous imaginons une chaîne de hautes montagnes, représentant » les hommes du XVIII^e siècle, terminée par une cime imposante » — Gauss —, puis une large et riche contrée remplie de nouveaux éléments de vie. » Ces nouveaux éléments de vie, ce sont les Cauchy, Abel, Galois, Bolzano (1781-1848), Jacobi, Riemann, Weierstrass, Cayley, Cantor, Kronecker, Dedekind, Lie, Hermite, Klein, Poincaré, Hilbert, Borel (1871-1956), Lebesgue (1875-1941), Elie Cartan (1869-1951), sans parler des mathématiciens encore vivants. Tous s'imposent sans conteste par l'ampleur et la profondeur de leurs vues.

CHRONOLOGIE DE L'HISTOIRE DES MATHÉMATIQUES

La présentation chronologique de l'histoire d'une science est sans aucun doute utile et éclairante, car elle en fait apparaître les moments majeurs, les étapes de son progrès, les synchronismes. Mais elle fausse toujours quelque peu la réalité du développement pour les raisons suivantes : telle œuvre, telle découverte qui se situent à un moment bien déterminé ont été précédées d'une préparation souvent longue, et elles sont suivies de développements, d'améliorations souvent très sensibles. De plus, une chronologie — et surtout celle-ci — ne peut retenir que les personnages, les œuvres et les découvertes les plus importants.

La présente chronologie ne débute qu'avec les mathématiques grecques et elle s'arrête à la fin du XIX^e siècle. Pour l'Antiquité préhellénique, une chronologie n'est pas possible, faute de données historiques suffisamment précises. Et, pour le XX^e siècle, les progrès sont à la fois trop nombreux et trop techniques pour prendre place dans un cadre qui peut ne consacrer que des vues sommaires aux travaux contemporains de la recherche mathématique.

Chronologie

	Géométrie	Arithmétique, Algèbre	Calcul infinitésimal, Intégration
VI^{e} siècle avant J.-C.	Thalès, fondateur de la géométrie		
V^{e} siècle avant J.-C.	Hippocrate de Chio : duplication du cube (construire un cube de volume double d'un cube donné) Théorème de Pythagore	Pythagore et son école : le monde est régi par les nombres, arithmo-géométrie	Pythagore : découverte des nombres irrationnels ($\sqrt{2}$)
IV^{e} siècle avant J.-C.	Platon : polyèdres réguliers Eudoxe : théorie des proportions Euclide : *Eléments* (partie géométrique) reste très longtemps l'ouvrage de référence	Platon : philosophie du nombre Euclide : Théorie élémentaire des entiers (livre VIII)	Euclide : nombres irrationnels (liv. VIII, X) théorie des grandeurs (liv. V)
III^{e} siècle avant J.-C.	Apollonius de Perge : coniques	Eratosthène : découverte du « crible » portant son nom pour la détermination des nombres premiers	Archimède : intégration par exhaustion ; calculs de surfaces, de volumes, (parabole, cône, sphère...), de centres de gravité
II^{e} siècle avant J.-C.			
I^{er} siècle avant J.-C.	Menelaus : relations dans les triangles (transversales)	Héron d'Alexandrie : méthodes de mesure, calculs approchés	
I^{er} siècle après J.-C.		Nicomaque de Gérase : *Introduction à l'arithmétique*	
II^{e} siècle après J.-C.	Ptolémée : début de la trigonométrie		
III^{e} siècle après J.-C.		Diophante : étude des équations du 1er et du 2^{e} degré en nombres entiers	
IV^{e} siècle après J.-C.	Pappus : *Collections mathématiques* (recueil de problèmes et de propositions)		
V^{e} siècle après J.-C.	Proclus : *Commentaires sur Euclide*		
VI^{e} siècle après J.-C.		Inde : invention du zéro	
VII^{e} siècle après J.-C.			
VIII^{e} siècle après J.-C.			
IX^{e} siècle après J.-C.		Al Khwarismi : premier traité d'algèbre Les Arabes empruntent le zéro à l'Inde	Thabit Ibn Qurra : calcule $\int \sqrt{x}$

	Géométrie	Arithmétique, Algèbre	Analyse, Calcul, Intégration
X^e siècle après J.-C.	Gerbert (le pape Sylvestre II)		
XI^e siècle après J.-C.		Al Kharki et Al Biruni : essai de résolution de l'équation du 3^e degré	Al Biruni : notions de vitesse instantanée et d'accélération Ibn Al Haytan : somme des puissances quatrièmes d'entiers successifs
XII^e siècle après J.-C.	Gérard de Crémone : traduction des mathématiciens arabes et, à travers eux, d'Euclide		
XIII^e siècle après J.-C.		Léonard de Pise (appelé Fibonacci) : introduction de l'algèbre des Arabes en Occident	Levi ben Gerson : analyse combinatoire (arrangements, permutations)
XIV^e siècle après J.-C.		Nicolas Oresme : nombres négatifs	
		Chuquet : utilisation des nombres négatifs et des exposants	
XVI^e siècle après J.-C.	Albert Dürer consacre un traité à la perspective	● Italie : Pacioli, Cardan, Tartaglia, Bombelli : création des nombres imaginaires, résolution des équations des 3^e et 4^e degrés. ● Allemagne : Rudolff, Stiefel ● France : Viète (introduit la représentation de l'inconnue par une lettre). ● Hollande : Stevin Emploi des symboles littéraux, calcul algébrique	Pendant tout le siècle, constitution progressive du calcul intégral et différentiel
1610	Neper : invention des logarithmes		
1620		Albert Girard : nombre des racines d'une équation algébrique	Grégoire de St-Vincent : notion de limite; calcul de l'aire de l'hyperbole $(\int_1^x)$; maniement des intégrales
1630	Descartes : création de la géométrie analytique; notion de tangente à une courbe Desargues : géométrie projective Pascal : *Essay pour les coniques*	Descartes : introduit les notations modernes en algèbre	Cavalieri : géométrie des indivisibles (pour le calcul d'aires) Fermat : étude des problèmes de minima et de maxima
1640			Sarasa : fonction logarithmique

	Géométrie	**Arithmétique, Algèbre**	**Analyse, Calcul, Intégration**
1650	Pascal : *Traité des coniques* (1668) et aire de la roulette (cycloïde)	Pascal : triangle arithmétique Fermat : $x^n + y^n = z^n$ n'a pas de solution en nombres entiers pour $n > 2$ (théorème non encore démontré)	Wallis : *Arithmetica infinitorum :* sommes et produits infinis, séries Pascal : intégration de fonctions polynômes, relations entre intégrales, triangle caractéristique (approche du calcul infinitésimal)
1660			
1670	Ph. de La Hire : *Sections coniques*		Leibniz et Newton posent les bases du calcul infinitésimal et algébrisent ces opérations (différentiation et intégration)
1680			Bernoulli : équations différentielles
1690			Rolle : théorème des accroissements finis L'Hospital *Analyse des infiniments petits* (premier traité systématique du calcul infinitésimal)
1700	Equation de la sphère		
1710			Taylor : développement en série d'une fonction utilisant ses dérivées successives
1720	Bernoulli et Euler : étude des lignes géodésiques	Euler : expression de sin x et cos x à l'aide des imaginaires	Riccati : résolution d'équations différentielles
1730	Clairaut : courbes gauches Saccheri : première tentative de géométrie non euclidienne (essai de démonstration par l'absurde de l'axiome des parallèles)		Euler : équations aux dérivées partielles
1740	Euler : surfaces à courbure constante; équations des surfaces du second degré	D'Alembert et Euler : théorème fondamental de l'algèbre, essai de démonstration (une équation algébrique de degré n a n racines) Cramer, Bezout : systèmes d'équations linéaires	Formule de Taylor - Maclaurin Bernoulli, Euler, d'Alembert : équation des cordes vibrantes Euler : *Introduction à l'analyse des infiniments petits*
1750			Euler : *Institutiones calculi integralis* Euler : notion de courbure d'une surface
1760	Lambert : progrès vers l'établissement de la géométrie non euclidienne	Lambert : trigonométrie sphérique	

	Géométrie	Arithmétique, Algèbre	Analyse, Calcul, Intégration
1770	Monge : équations des surfaces algébriques	Lagrange : relations entre les fonctions symétriques des racines et les coefficients d'une équation algébrique	
1780	Lagrange : équations de la droite et du plan; emploi des 3 axes de coordonnées		
1790	Legendre : *Eléments de géométrie* Monge : *Traité de géométrie descriptive*	Legendre : *Théorie des nombres* Gauss : démonstration complète du théorème fondamental de l'algèbre	Lagrange : *Leçons sur la théorie des fonctions*
1800		Gauss : *Disquisitiones arithmeticae,* ouvrage fondamental créant la théorie des groupes abéliens et résolvant de nombreux problèmes d'arithmétique Argand : représentation géométrique des nombres imaginaires	Fourier : Représentation des fonctions par des séries trigonométriques
1810			Bolzano : premières définitions rigoureuses des notions de limite et de continuité
1820	Poncelet : *Traité des propriétés projectives des figures* Plücker : coordonnées homogènes Lobatchevsky : géométrie non euclidienne		Cauchy : *Cours d'analyse :* principes de la théorie des séries, étude des fonctions continues, rigueur dans la notion d'intégrale Cauchy : bases de l'étude des fonctions d'une variable complexe (fonctions analytiques) Jacobi : étude des fonctions elliptiques
1830	Bolyai : géométrie non euclidienne	Galois : théorie définitive des équations algébriques fondée sur la notion de groupe et celle de corps Gauss : théorie des nombres complexes Bellavitis : concept de vecteur	
1840		Grassmann : produit intérieur et scalaire de deux vecteurs Liouville : première démonstration de la transcendance d'un nombre Cayley : théorie des matrices Boole, de Morgan : algèbre de la logique	Bolzano : construction d'une fonction continue n'ayant de dérivée en aucun point Weierstrass : notion de convergence uniforme

	Géométrie	**Arithmétique, Algèbre**	**Analyse, Calcul, Intégration**
1850	Riemann : surface de Riemann, début du développement de la géométrie algébrique Laguerre : définition projective de la distance Riemann : *Sur les hypothèses qui servent de base à la géométrie* Riemann : *Analysis situs* (création de la topologie) Cayley : théorie générale des géométries euclidiennes et non euclidiennes	Hamilton : création des quaternions (premier exemple d'un corps non commutatif)	Riemann : étude des fonctions d'une variable complexe Riemann généralise la notion d'intégrale introduite par Cauchy
1860		Jordan : groupes de substitutions Christoffel : calcul tensoriel	Weierstrass : étude des fonctions d'une variable complexe
1870	Lie : groupes de transformations continues Klein : Programme d'Erlangen (synthèse des géométries, fondée sur la notion de groupe de transformation)	Dedekind : notion d'idéal d'un anneau Kronecker : théorie des corps de nombres algébriques	Weierstrass : étude des fonctions non dérivables Cantor et Dedekind : définition arithmétique des nombres irrationnels Heine : notion de continuité uniforme Cantor : notion d'ensemble, de puissance d'un ensemble
1880		Lindemann : transcendance du nombre Ricci : calcul tensoriel Peano : espace vectoriel Dedekind : définition des entiers sur la base de la théorie des ensembles Peano : axiomatique du nombre et de la géométrie	Poincaré : fonctions fuchsiennes Cantor : ensembles bien ordonnés Cantor : nombres ordinaux
1890	Hilbert : Fondements de la géométrie première théorie mathématique en forme axiomatique moderne	Levi Civita : calcul différentiel absolu Premiers paradoxes de la théorie des ensembles (non encore axiomatisée) : ensemble de tous les ensembles	Intégrale de Stieltjes
1900	Hausdorff : notion de voisinage, définition axiomatique des espaces topologiques Fréchet : notion d'espace métrique		Intégrale de Lebesgue Zermelo prouve l'existence d'un bon ordre sur tout ensemble Zermelo : première axiomatisation de la théorie des ensembles (introduction de l'axiome du choix) Radon : définition d'une mesure
1910			

HOMÉOMORPHISME, HOMÉOMORPHE ◆ page 529.

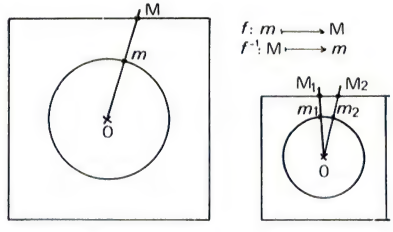
◆ Mot du vocabulaire de topologie. Lire d'abord Topologie, Application et Bijection.

Exemple

Dans le plan, considérons un cercle et un carré de même centre O. Toute demi-droite issue de O coupe le cercle en m et le carré en M. L'application f qui à m fait correspondre M est telle que :

— à chaque point de l'une des deux figures correspond un point de l'autre ;

— à deux points voisins de l'une des deux figures correspondent deux points voisins de l'autre.



Cette application f est un homéomorphisme.

D'une façon générale, une application f d'un espace topologique E vers un espace topologique F est un homéomorphisme si elle est bijective, continue et si son application réciproque f^{-1} est aussi continue (on exprime souvent cela en disant plus brièvement que f est bijective et bicontinue).

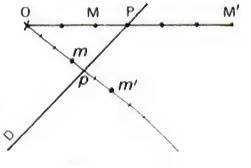
L'image par f d'un ouvert quelconque de E est un ouvert de F et l'image par f^{-1} d'un ouvert de F est un ouvert de E.

S'il existe un homéomorphisme de l'espace E sur l'espace F, on dit que les deux espaces topologiques E et F sont homéomorphes.

HOMOGRAPHIQUE (Fonction) : voir Fonction et les pages 158, 114, 312.

HOMOLOGIE ◆

◆ Lire d'abord Birapport.



Soient O un point, D une droite et k un nombre réel. On appelle homologie de centre O, de directrice D et de rapport k l'application qui à chaque point M du plan associe un point M′ (unique) tel que :

— O, M et M′ soient alignés ;

— et, si P est le point d'intersection de la droite OMM′ avec (D), le birapport des quatre points O, P, M, M′ soit égal à k.

Dans cette application, le point O et la droite (D) sont invariants.

Cas où $k = -1$: homologie harmonique.

◆ La définition donnée ci-dessous faisant intervenir un grand nombre de notions de la théorie des ensembles et de l'algèbre, il est conseillé au lecteur non familiarisé avec ces notions de lire les articles « les Ensembles » et « l'Algèbre ».

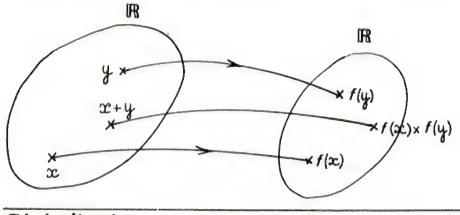
HOMOMORPHISME ou MORPHISME ◆

Considérons l'application de $\mathbb{R}$ (ensemble des nombres réels) dans $\mathbb{R}$ qui à tout nombre réel x fait correspondre le nombre 2^x (2 puissance x).

L'image du nombre x est 2^x ; l'image d'un autre nombre y est 2^y ; l'image du nombre $x + y$ est 2^{x+y} ; or : $2^{x+y} = 2^x \times 2^y$◆, c'est-à-dire que l'image de la

◆ Voir Exponentielle.

somme $x + y$ est égale au *produit* des deux images 2^x et 2^y. On dit que l'application ci-dessus est un homomorphisme de l'ensemble $\mathbb{R}$ muni de l'addition, dans l'ensemble $\mathbb{R}$ muni de la multiplication.



La fonction logarithme est aussi un homomorphisme, mais de $\mathbb{R}_+^*$ (ensemble des réels strictement positifs) muni de la multiplication dans $\mathbb{R}$ muni de l'addition, car : $\log(x \times y) = \log x + \log y$.

Généralisation

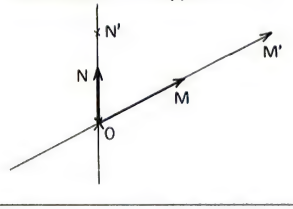
Soient E un ensemble muni d'une loi de composition interne notée T et F un ensemble muni d'une loi de composition interne notée $\perp$; une application de E dans F est un morphisme si, quels que soient x et y, l'image par f du composé $x \mathrm{T} y$ est le composé des images $f(x) \perp f(y)$:

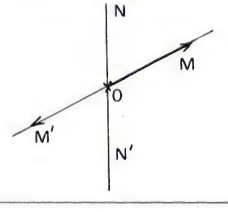
$$f(x \mathrm{T} y) = f(x) \perp f(y).$$

Un morphisme f de l'ensemble E muni de la loi de composition T (E, T) dans l'ensemble F muni de la loi de composition $\perp$ (F, $\perp$) est dit :
— monomorphisme, si l'application f est injective ;
— épimorphisme, si l'application f est surjective ;
— isomorphisme, si l'application f est bijective ;
— endomorphisme, si les deux ensembles E et F sont égaux ;
— automorphisme, si E = F et f isomorphisme.

HOMOTHÉTIE pages 76 et 77, 264, 453.

Soient O un point du plan (ou de l'espace) et k un nombre réel non nul ; l'homothétie de centre O et de rapport k est la transformation qui à chaque point M du plan (ou de l'espace) associe le point M′ tel que le vecteur $\overrightarrow{OM'}$ soit égal au produit par k du vecteur $\overrightarrow{OM}$: $\overrightarrow{OM'} = k \cdot \overrightarrow{OM}$.





Si $k = 1$, l'homothétie est la transformation identique (car tout point M se transforme alors en lui-même).
La composée de deux homothéties de centre O et de rapports respectifs k et k' est l'homothétie de centre O et de rapport $k \times k'$.
L'ensemble des homothéties de centre O, muni de la loi de composition des applications, est un groupe commutatif◆.

◆ Voir l'article « l'Algèbre », page 29.

HYPERBOLE : voir Coniques et les pages 225, 397.

HYPERBOLOÏDE

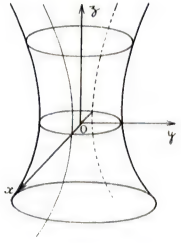
Surface dont l'équation est du second degré, qui admet un centre de symétrie, trois axes de symétrie qui passent par ce centre et dont certaines sections planes sont des hyperboles. Il y a deux sortes d'hyperboloïdes :

Hyperboloïde à une nappe

Dans un système d'axes rectangulaires qui sont les axes de symétrie de la figure, l'équation d'un hyperboloïde à une nappe est :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} - 1 = 0.$$

Les sections par des plans parallèles au plan xOy sont des ellipses et les sections par des plans parallèles au plan yOz sont des hyperboles.



Si les nombres a et b sont égaux, l'hyperboloïde est un hyperboloïde de révolution ; les sections par des plans parallèles à xOy sont des cercles. Un tel hyperboloïde peut être engendré par la rotation d'une hyperbole autour de son axe non transverse.

Hyperboloïde à deux nappes

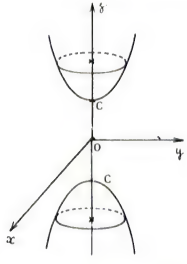
Dans un système d'axes rectangulaires qui sont les axes de symétrie de la figure, l'équation d'un hyperboloïde à deux nappes est :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0.$$

Les sections par des plans parallèles à zOx et zOy sont des hyperboles et les sections par des plans parallèles à xOy sont vides tant que :

$$1 - \frac{z^2}{c^2} > 0 \quad (-c < z < +c).$$

Si $z > c$ ou $z < -c$, alors ces sections sont des ellipses.



Si $a = b$, l'hyperboloïde est de révolution ; les sections par des plans parallèles à xOy, quand elles existent, sont des cercles. Un tel hyperboloïde peut être engendré par la rotation d'une hyperbole autour de son axe transverse.

HYPERGÉOMÉTRIQUE (Loi) : voir l'article « les Statistiques », pages 490 et 491.

I

ICOSAÈDRE : voir Polyèdres

IDEMPOTENT ◆

◆ Voir les articles « les Ensembles » et « l'Algèbre ».

Un élément x d'un ensemble E muni d'une loi de composition interne (notée $\divideontimes$) est idempotent si, composant x avec lui-même, on trouve x comme résultat :

$$x \divideontimes x = x.$$

Dans l'ensemble des entiers naturels, muni de la multiplication, les seuls éléments idempotents sont 0 et 1 :

$$0 \times 0 = 0 \qquad 1 \times 1 = 1.$$

Dans l'ensemble $\mathfrak{P}(E)$ des parties d'un ensemble E, muni de la loi intersection, tout élément est idempotent, car, pour toute partie A de E, l'intersection de A avec elle-même est égale à A : $A \cap A = A$.
De même pour $\mathfrak{P}(E)$ muni de la loi réunion, tout élément est idempotent, car, pour tout A de $\mathfrak{P}(E)$, on a :

$$A \cup A = A.$$

Si tout élément de E est idempotent, la loi $\divideontimes$ est dite idempotente et E lui-même est dit idempotent.

IDENTIQUE, IDENTITÉ pages 245, 253, 282.

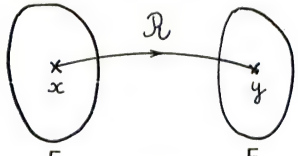
Si E est un ensemble, l'application de E dans E qui à tout élément x de E fait correspondre x lui-même est appelée application identique de E (ou identité de E) et se note Id_E : $Id_E : x \longrightarrow x$.

Fonction identité : voir Fonction et les pages 114, 125, 312.

IDENTITÉ REMARQUABLE : voir Binôme

IMAGE◆ pages 76, 79, 114, 122, 200, 274, 277, 445.

◆ Lire d'abord Relation, puis Application.



$x \mathcal{R} y$ y image de x par $\mathcal{R}$

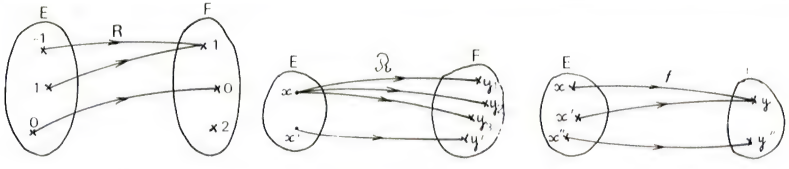
Soient E et F deux ensembles et $\mathcal{R}$ une relation de E vers F. Si l'élément x de E est en relation, par $\mathcal{R}$, avec l'élément y de F, on dit que y est image de x par $\mathcal{R}$ et que x est antécédent de y par $\mathcal{R}$.

Exemple : $E = \{ -1, 0, 1 \}$ $F = \{ 0, 1, 2 \}$

La relation « a pour carré » définie de E vers F est telle que :

— 1 a pour image 1	0 a pour antécédent 0
1 a pour image 1	1 a deux antécédents : — 1 et 1
0 a pour image 0	2 n'a pas d'antécédent

Dans le cas d'une relation fonctionnelle, chaque élément x de E a au plus une image. Dans le cas d'une application, chaque élément de E a une seule image.



Relation R « a pour carré » Relation de E vers F Application

Image directe d'un sous-ensemble

Soient E et F deux ensembles, A une partie de E et f une application de E vers F ; l'ensemble constitué des éléments de F qui sont images d'au moins un élément de A est appelé image directe (ou image) de A par l'application f et noté $f(A)$:

$$f(A) = \{ y \mid y \in F, \exists x \in A, y = f(x) \}$$

(lire : f de A est l'ensemble des y appartenant à F tels qu'il existe au moins un x appartenant à A tel que y soit l'image de x).

Exemple :

Si f est l'application de l'ensemble des entiers naturels $\mathbb{N}$ dans $\mathbb{N}$, qui à tout nombre entier naturel fait correspondre son carré et si A est la partie de $\mathbb{N}$ égale à $\{ 1, 2, 3 \}$ l'image directe de A est :

$$f(A) = \{ 1, 4, 9 \}$$

— Il est clair que $f(E)$ est incluse dans F (elle est égale à F si l'application f est surjective).

— Si A et B sont deux parties de E telles que A soit incluse dans B, alors l'image de A est incluse dans l'image de B :

$$A \subset B \Longrightarrow f(A) \subset f(B)$$

— L'image de la réunion de deux parties de E est égale à la réunion des images de ces deux parties :

$$f(A \cup B) = f(A) \cup f(B)$$

— Mais l'image de l'intersection de deux parties de E est seulement incluse dans l'intersection des images (il y a égalité dans le cas où f est injective) :

$$f(A \cap B) \subset f(A) \cap f(B)$$

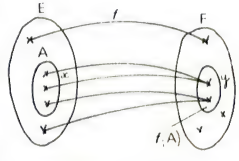


Image directe de A.

Image réciproque d'un sous-ensemble

Soient E et F deux ensembles, A' une partie de F et f une application de E vers F.

Le sous-ensemble de E constitué des éléments de E, dont l'image par f appartient à A′, est appelé image réciproque de A′ et notée $f^{-1}(A')$:

$$f^{-1}(A') = \{ x \mid x \in E, f(x) \in A' \}$$

(lire : f — 1 de A′ est l'ensemble des x appartenant à E tels que f de x appartienne à A′).

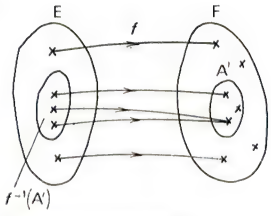


Image réciproque de A′.

Exemple :

Pour la même application que celle de l'image directe : si $A' = \{ 1, 2, 3, 4 \}$, l'image réciproque de A′ est :

$$f^{-1}(A') = \{ 1, 2 \}$$

(car 1 et 2 sont les deux seuls entiers dont les carrés appartiennent à A′).

— L'image réciproque de F est évidemment E (puisque les images de tous les éléments de E appartiennent à F).

— L'image réciproque de l'ensemble vide est vide, mais il peut arriver que l'image réciproque d'un sous-ensemble non vide de F soit vide (il suffit que ce sous-ensemble de F soit formé d'éléments n'ayant pas d'antécédent dans E).
Par exemple, si $A'' = \{ 2, 3 \}$, l'image réciproque de A″ est vide, car aucun entier n'a 2 pour carré et aucun entier n'a 3 pour carré.

— Si A′ et B′ sont deux parties de F telles que A′ soit incluse dans B′, alors l'image réciproque de A′ est incluse dans l'image réciproque de B′ :

$$A' \subset B' \Longrightarrow f^{-1}(A') \subset f^{-1}(B')$$

— L'image réciproque de la réunion de deux parties de F est la réunion des images réciproques de ces deux parties :

$$f^{-1}(A' \cup B') = f^{-1}(A') \cup f^{-1}(B')$$

— L'image réciproque de l'intersection de deux parties de F est l'intersection des images réciproques de ces deux parties :

$$f^{-1}(A' \cap B') = f^{-1}(A') \cap f^{-1}(B').$$

IMAGINAIRE (Nombre) pages 159 et 160, 186, 229, 327, 328, 329, 375.

Nombre imaginaire et nombre complexe sont deux expressions synonymes. Un nombre complexe est qualifié d'« imaginaire pur » s'il est de la forme $b \times i$, où b est un nombre réel et i un des deux nombres complexes (l'autre étant $-i$), dont le carré i^2 est négatif et égal à — 1.

IMPAIR◆

◆ Lire d'abord Application.

Une application f d'une partie P◆ de $\mathbb{R}$, ensemble des nombres réels, dans $\mathbb{R}$ est impaire si, pour tout x de P, on a :

◆ P étant telle que, si x lui appartient, alors $-x$ lui appartient aussi.

$$f(-x) = -f(x).$$

Exemple :

les fonctions de $\mathbb{R}$ dans $\mathbb{R}$: sinus, tangente, sont des fonctions impaires.

IMPLICATION

p	q	$p \Longrightarrow q$
V	V	V
V	F	F
F	V	V
F	F	V

Si p et q sont deux propositions, la proposition « $p \Longrightarrow q$ » (qui se lit : p implique q), nommée implication de q par p, est la proposition qui est fausse uniquement dans le cas où p est vraie et q est fausse. Sa table de vérité est représentée ci-contre.

IMPLICITE (Fonction)

Si f est une fonction réelle des deux variables réelles x et y telle que :

$$f : (x, y) \longmapsto f(x, y),$$

et si, pour chaque valeur x d'un sous-ensemble E de $\mathbb{R}$ (ensemble des nombres réels), l'équation $f(x, y) = 0$ a une seule solution en y, l'application (de E dans $\mathbb{R}$) qui à chaque valeur x de E fait correspondre la solution y correspondante est appelée fonction implicite.

Par exemple, soit $f(x, y) = x^2 + y^3 - 1$. Pour chaque valeur x de l'intervalle fermé [0, 1], l'équation en y : $y^3 + x^2 - 1 = 0$ (ou $y^3 = 1 - x^2$) a une seule solution :

$$y = \sqrt[3]{1 - x^2} \text{ (racine cubique de } 1 - x^2).$$

En général, on ne sait pas calculer l'expression de y en fonction de x ; on peut cependant effectuer certains calculs ; par exemple, si la fonction f admet des dérivées partielles $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ par rapport à x et à y, pour toute valeur de x pour laquelle la dérivée partielle par rapport à y n'est pas nulle la dérivée y' (de la fonction implicite par rapport à x) est égale à :

$$y' = -\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}.$$

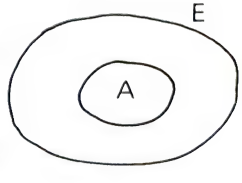
INCERTITUDE : voir Erreur

INCLUS pages 259, 309.

Un ensemble A est inclus dans un ensemble E si et seulement si tout élément de A est élément de E ; on note $A \subset E$:

$$(A \subset E) \Longleftrightarrow (\forall\, x \in A \Longrightarrow x \in E)$$

(lire : A inclus dans E est équivalent à : quel que soit x appartenant à A, x appartient à E).



Exemple : $E = \{1, 2, 3, 4, 5\}$ $A = \{1, 2, 5\}$ $B = \{1, 2, 6\}$

A est inclus dans E, mais B ne l'est pas (car 6 est élément de B sans être élément de E).

Quels que soient les ensembles A, B, C, on a :

— $A \subset A$; A est inclus dans lui-même.

— Les deux ensembles A et B sont égaux si et seulement si A est inclus dans B et B est inclus dans A : $(A = B) \Longleftrightarrow (A \subset B \text{ et } B \subset A)$.

— Si A est inclus dans B et si B est inclus dans C, alors A est inclus dans C :
$(A \subset B \text{ et } B \subset C) \Longrightarrow (A \subset C)$.

INDÉPENDANCE LINÉAIRE : voir, dans l'article « l'Algèbre linéaire », Vecteurs libres, page 67.

INDICE

Pour repérer les termes d'une suite (ou, plus généralement, les éléments d'un ensemble dénombrable), on les note u_1 (on lit : u un, ou : u indice un), u_2 ... ; les nombres 1, 2, ..., sont les indices.

Indice muet

Dans les écritures :

$$\sum_{i=1}^{n} x_i, \quad \bigcap_{i=1}^{n} E_i, \quad \bigcup_{i=1}^{n} E_i,$$

la lettre i est un indice muet : elle doit prendre successivement les valeurs 1, 2, 3, ..., n. Ainsi, $\sum_{i=1}^{n} x_i$ signifie que l'on doit faire la somme de tous les x_i, i variant de 1 jusqu'à n, c'est-à-dire :

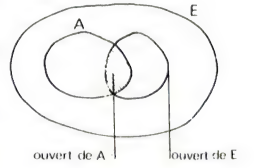
$$x_1 + x_2 + x_3 + \ldots + x_n.$$

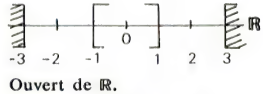
INDUITE (Topologie) (Sous-espace topologique)◆

◆ Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

Si E est un espace topologique et A une partie de E, on appelle topologie induite sur A par la topologie de E la topologie qui admet pour ouverts les intersections de A avec les ouverts de E.
L'ensemble A muni de la topologie induite par celle de E est appelé sous-espace topologique de E.
La topologie induite sur A n'a pas forcément les mêmes propriétés que la topologie de E.

Exemple : $\mathbb{Z}$ (ensemble des nombres entiers relatifs négatifs et positifs) est un sous-ensemble de $\mathbb{R}$ (ensemble des nombres réels). Si $\mathbb{R}$ est muni de la topologie de l'ordre, la topologie induite sur $\mathbb{Z}$ est la topologie discrète (fine) (car un ouvert de $\mathbb{R}$ pour la topologie de l'ordre est une réunion d'intervalles ouverts de $\mathbb{R}$; l'intersection d'une telle réunion avec $\mathbb{Z}$ est une partie de $\mathbb{Z}$; toutes les parties de $\mathbb{Z}$ peuvent être obtenues comme intersection avec $\mathbb{Z}$ d'un ouvert de $\mathbb{R}$, donc toute partie de $\mathbb{Z}$ est un ouvert).





Ouvert de $\mathbb{R}$.
$]-3, -1[\cup]1, 3[$ est un ouvert de $\mathbb{R}$.
L'intersection de cet ouvert avec $\mathbb{Z}$ est $\{-2, 2\}$ qui est une partie de $\mathbb{Z}$.

INFÉRIEUR◆

Si E est un ensemble muni d'une relation d'ordre total notée $\mathcal{R}$, si l'élément x de l'ensemble E est en relation avec y de E, au lieu de noter $x \mathcal{R} y$ (x relation y), on note $x \leqslant y$ (on lit : x inférieur ou égal à y). On dit aussi que y est supérieur ou égal à x.
Si l'égalité $x = y$ ne peut avoir lieu, on note $x < y$ (on lit : x strictement inférieur à y, ou : y strictement supérieur à x).

◆ Lire d'abord Ordre.

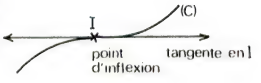
INFINI (Ensemble) : voir l'article « les Ensembles », page 265, l'article « les Nombres », pages 375 à 377, et les pages 105, 299, 418.

INFINITÉSIMAL (Calcul) : voir l'article « l'Analyse », page 125 et suiv. et les pages 326 à 330.

INFLEXION (Point d')

Point où la concavité d'une courbe plane change de sens ou encore point en lequel une courbe traverse sa tangente.

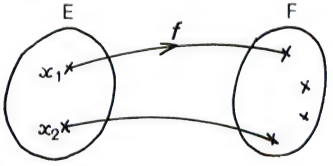
Si l'équation cartésienne de la courbe (C) est $y = f(x)$ (où f est une fonction deux fois dérivable), une condition nécessaire et suffisante pour qu'il y ait un point d'inflexion est que la dérivée seconde f'' s'annule en changeant de signe (si elle s'annule sans changer de signe, il n'y a pas nécessairement inflexion).



INJECTION◆ pages 80, 206, 211, 277 et 278, 335.

◆ Lire d'abord Application.

Soient deux ensembles E et F et une application f de E vers F.



Si deux éléments quelconques, distincts, de E, x_1 et x_2, ont pour images, par l'application f, deux éléments distincts dans F, $f(x_1)$ et $f(x_2)$, on dit qu'il y a injection ou que l'application est injective.

On note :

$$\forall\, x_1 \in \mathrm{E},\quad \forall\, x_2 \in \mathrm{E},\qquad x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2)$$

(lire : quels que soient x_1 et x_2 appartenant à l'ensemble E, x_1 étant distinct de x_2, l'image de x_1 par l'application f : $f(x_1)$, est distincte de l'image de x_2 : $f(x_2)$).

Si f est une application injective de E vers F, tout élément de F est image d'au plus un élément de E, donc deux images $f(x_1)$ et $f(x_2)$ ne peuvent être égales que si x_1 est égal à x_2 : $f(x_1) = f(x_2) \Longrightarrow x_1 = x_2$).

Si E et F sont des ensembles finis, le nombre d'éléments de F est donc supérieur ou égal au nombre d'éléments de E.

Exemples :

— *d'injection* dans l'ensemble des nombres réels $\mathbb{R}$: $f : x \longmapsto 2x$

— *de non-injection :* $x \longmapsto x^2$ car $(-1)^2 = (1)^2$ et $-1 \neq 1$.

INSCRIPTIBLE : voir Circonscrit

INTÉGRALE pages 134 à 147, 151, 165, 177, 180, 292, 429.

Calcul de l'aire limitée par une courbe et l'axe Ox sur un intervalle $[a, b]$: voir l'article « l'Analyse », page 134.

Intégrale définie et indéfinie

Toutes les primitives G d'une fonction f peuvent être calculées à partir de l'une d'entre elles F par la formule :

$$G(x) = F(x) + C,$$

où C est une constante quelconque. On note :

$$G(x) = \int f(x)\,dx,$$

et on l'appelle intégrale indéfinie, par opposition à $\int_a^b f(x)\,dx$, qu'on appelle intégrale définie (dans l'intervalle (a, b)).

INTÉRIEUR◆

◆ Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

Si A est une partie d'un espace topologique E, la réunion de tous les ouverts de E contenus dans A est un ouvert de E qui est contenu dans A (c'est le plus grand de tous les ouverts de E contenus dans A, puisqu'il les contient tous) ; on appelle cet ouvert intérieur de A et on le note $\mathring{A}$ ◆.

◆ Extérieur de A = intérieur du complémentaire de A $\Longleftrightarrow$ intérieur de A = extérieur du complémentaire de A.

Exemple :

Soit un intervalle quelconque (a, b) $(a < b)$ de la droite numérique $\mathbb{R}$ (ensemble des nombres réels), l'intérieur de l'intervalle (a, b) est l'intervalle ouvert $]a, b[$.

Propriétés

— Si A est un ouvert, l'intérieur de A est égal à A, et réciproquement :

$$A \text{ ouvert} \Longleftrightarrow A = \mathring{A}.$$

— Si A et B sont deux parties du même espace topologique E et si A est contenue dans B, alors l'intérieur de A est contenu dans l'intérieur de B :

$$A \subset B \Longrightarrow \mathring{A} \subset \mathring{B}.$$

— L'intérieur de l'intersection de deux parties A et B de E est égal à l'intersection de l'intérieur de A et de l'intérieur de B :

$$\overset{\circ}{\widehat{A \cap B}} = \mathring{A} \cap \mathring{B}.$$

Mais la réunion A ∪ B n'a pas pour intérieur la réunion des intérieurs de A et de B ; la réunion des intérieurs de A et de B est contenue dans l'intérieur de la réunion A ∪ B :

$$\mathring{A} \cup \mathring{B} \subset \overset{\circ}{\widehat{A \cup B}}.$$

Ainsi, si E est l'ensemble $\mathbb{R}$ des nombres réels, A l'ensemble des nombres rationnels $\mathbb{Q}$, et B l'ensemble des nombres irrationnels $\mathbb{R} - \mathbb{Q}$, la réunion de A et de B est l'ensemble $\mathbb{R}$; l'intérieur de cette réunion est donc $\mathbb{R}$. Mais si l'intérieur de A est vide, l'intérieur de B est vide aussi ◆, donc la réunion de l'intérieur de A et de l'intérieur de B est vide :

$$\overset{\circ}{\widehat{A \cup B}} = \mathbb{R}$$

$$\mathring{A} \cup \mathring{B} = \varnothing.$$

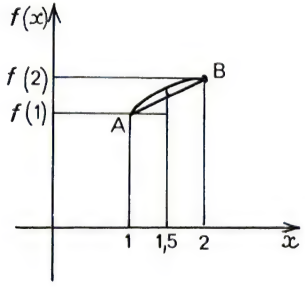
◆ En effet, aucun ouvert non vide de $\mathbb{R}$ ne peut être contenu dans $\mathbb{Q}$ (un ouvert non vide de $\mathbb{R}$ contient à la fois des nombres rationnels et des nombres irrationnels et $\mathbb{Q}$ ne contient que des rationnels) : aucun ouvert non vide de $\mathbb{R}$ ne peut être contenu dans $\mathbb{R} - \mathbb{Q}$ (car $\mathbb{R} - \mathbb{Q}$ ne contient que des irrationnels). Le seul ouvert qui puisse être contenu dans $\mathbb{Q}$, comme dans $\mathbb{R} - \mathbb{Q}$, est l'ensemble vide.

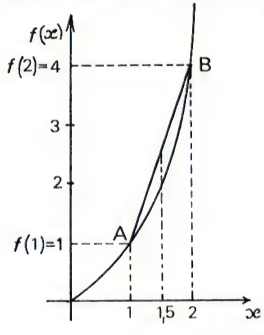
INTERNE (Loi de composition) : voir Composition (Loi de) et les pages 57, 60, 61.

INTERPOLATION LINÉAIRE

Supposons vouloir calculer la valeur d'une fonction réelle f d'une variable réelle x pour la valeur $x = 1{,}5$ et qu'on connaisse la valeur de f pour 1 et la valeur de f pour 2 ; par exemple, $f(1) = 3$ et $f(2) = 5$. Si on n'est pas capable de calculer $f(1{,}5)$ directement, on fait l'hypothèse simplificatrice suivante ◆ : entre 1 et 2, l'accroissement de la fonction est proportionnel à l'accroissement de la variable.

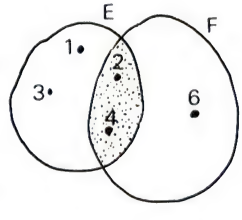
◆ Cela consiste, en fait, à remplacer entre A et B l'arc de la courbe représentative de f par la droite AB.





Si $f(x)$ augmente de 2 quand x augmente de 1, quand x augmentera de 0,5 (la moitié de 1), $f(x)$ augmentera de la moitié de 2, c'est-à-dire de 1. La valeur ainsi trouvée de $f(1{,}5)$ est $3 + 1 = 4$; c'est une valeur approchée de $f(1{,}5)$.
Pour la fonction $f(x) = x^2$, on a $f(1) = 1$ et $f(2) = 4$; le calcul direct donne pour $(1{,}5)^2$: 2,25. L'interpolation linéaire donne $1 + \frac{3}{2} = 2{,}5$.

INTERSECTION pages 22, 23, 25, 26, 190, 249, 265, 309, 363, 364.



— Si E et F sont deux ensembles, on appelle intersection de E et de F et on note $E \cap F$ (lire : E inter F) l'ensemble des éléments communs aux deux ensembles E et F :

$$E \cap F = \{ x \mid x \in E \text{ et } x \in F \}$$

(lire : E inter F est égal à l'ensemble des x tels que x appartienne à E et x appartienne à F).

— Quels que soient les ensembles E et F, l'intersection de E et de F est égale à l'intersection de F et de E :

$$E \cap F = F \cap E.$$

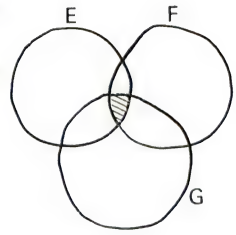
— Quel que soit E, l'intersection de E avec l'ensemble vide est vide :

$$E \cap \varnothing = \varnothing.$$

— Si on introduit un troisième ensemble G, quels que soient les ensembles E, F et G, l'intersection de E et de l'intersection de F et de G est égale à l'intersection de l'intersection de E et de F avec G :

$$E \cap (F \cap G) = (E \cap F) \cap G.$$

On appelle cet ensemble l'intersection des trois ensembles E, F et G et on le note (sans parenthèses) : $E \cap F \cap G$.



$E \cap F \cap G$

— Si $E_1, E_2, \ldots, E_n$ sont n ensembles, l'intersection de ces n ensembles (ensemble des éléments communs à tous ces ensembles) est notée $\bigcap_{i=1}^{n} E_i$.

INTERVALLE pages 102, 111, 112, 122, 130, 134 à 147, 239, 205, 230, 239, 311.

Soit $\mathbb{R}$, l'ensemble des nombres réels, ordonné par la relation « inférieur ou égal » et soient deux nombres réels a et b tels que a soit strictement inférieur à b. On appelle :

— *Intervalle ouvert* d'origine a et d'extrémité b et on note $]a, b[$ l'ensemble des nombres réels x strictement compris entre a et b :

$$x \in]a, b[\iff a < x < b$$

(lire : x appartient à l'intervalle ouvert a, b si et seulement si x est strictement compris entre a et b).

— *Intervalle fermé* (ou segment) d'origine a et d'extrémité b et on note $[a, b]$ l'ensemble des nombres réels x supérieurs ou égaux à a et inférieurs ou égaux à b :

$$x \in [a, b] \iff a \leqslant x \leqslant b.$$

— *Intervalle semi-ouvert* à droite (ou à gauche) $[a, b[$ (ou $]a, b]$) l'ensemble des nombres réels x supérieurs ou égaux à a et strictement inférieurs à b (ou strictement supérieurs à a et inférieurs ou égaux à b) :

$$x \in [a, b[\iff a \leqslant x < b$$
$$x \in]a, b] \iff a < x \leqslant b.$$

— *Intervalle fermé (ou ouvert) illimité à gauche*, d'extrémité b, l'ensemble des nombres réels x inférieurs ou égaux (ou strictement inférieurs) à b et on le note $]-\infty, b]$ (ou $]-\infty, b[$) :

$$x \in]-\infty, b] \iff x \leqslant b$$
$$x \in]-\infty, b[\iff x < b.$$

— *Intervalle fermé (ou ouvert) illimité à droite*, d'origine a, l'ensemble des nombres réels x supérieurs ou égaux (ou strictement supérieurs) à a et on le note $[a, +\infty[$ (ou $]a, +\infty[$) :

$$x \in [a, +\infty[\iff x \geqslant a$$
$$x \in]a, +\infty[\iff x > a.$$

$\mathbb{R}$ lui-même peut être écrit $]-\infty, +\infty[$:

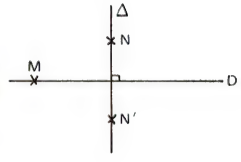
$$x \in]-\infty, +\infty[\iff x \in \mathbb{R}.$$

Ces définitions d'intervalles ouvert, fermé ou semi-ouvert peuvent être étendues à n'importe quel ensemble E, pourvu qu'il soit muni d'une relation d'ordre total.

Intervalles emboîtés : voir l'article « l'Analyse », page 102.

INVARIANT page 529.

Soit, dans un plan, la symétrie par rapport à une droite D.



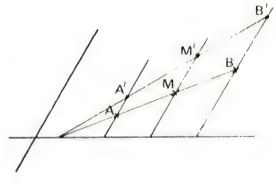
— Tout point M de la droite D se transforme en lui-même dans cette symétrie.

— Si Δ est une droite perpendiculaire à D, tout point N de Δ se transforme dans cette symétrie en un point N′ qui appartient encore à Δ, mais qui est différent de N (sauf si N appartient à la fois à D et à Δ).

La droite D est dite invariante point par point dans la symétrie par rapport à D ; la droite Δ est dite globalement invariante dans cette symétrie.

D'une façon générale, une figure F est invariante point par point par une transformation T si chaque point de F est transformé en lui-même par T ; elle est invariante globalement si chaque point de F est transformé en un (autre) point de F par T.

Invariant pour une famille de transformations



Soit E un ensemble sur lequel est définie une famille de transformations. On dit qu'une propriété d'un ensemble F d'éléments de E est invariante par cette famille (ou est un invariant de cette famille) si l'ensemble F′ transformé de F par n'importe quelle transformation de la famille a la même propriété que F.

Ainsi, dans le plan, la propriété, pour le point M, d'être au milieu du segment AB est une propriété invariante par n'importe quelle affinité.

INVERSE : voir Symétrique

INVERSE (Matrice)

Si M est une matrice carrée d'ordre n à coefficients réels, une matrice M′ (carrée, d'ordre n, à coefficients réels) est dite inverse de la matrice M si le produit de ces deux matrices est égal à la matrice unité I, avec $I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$:

$$M \times M' = I.$$

Si une telle matrice M′ existe, on a alors : $M' \times M = I$◆.

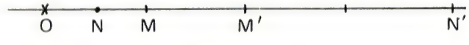
◆ Bien que la multiplication des matrices ne soit pas commutative, dans le cas de matrices inverses, $M \times M' = M' \times M$.

Si M′ est inverse de M, M est inverse de M′ ; on dit simplement que M et M′ sont deux matrices inverses.

Une matrice M n'admet pas toujours de matrice inverse. Dans le cas où elle en admet une, on dit qu'elle est inversible. Une condition nécessaire et suffisante pour qu'une matrice soit inversible est que son déterminant soit différent de 0.

INVERSION page 47.

Étant donné un point O, un nombre réel k et une unité de longueur, on appelle inversion de pôle O et de puissance k la transformation géométrique qui à tout point M du plan (ou de l'espace) fait correspondre le point M′ tel que :



Inversion de puissance 2.

— les points O, M et M′ soient alignés ;

— le produit des mesures algébriques de OM et de OM′ soit égal à k :

$$\overline{OM} \cdot \overline{OM'} = k.$$

$\overline{OM} \cdot \overline{OM'} = \overline{ON} \cdot \overline{ON'} = 2$.

En géométrie plane, une droite ne passant pas par O est transformée (dans une inversion de pôle O) en un cercle passant par O, tandis qu'une droite passant par O est globalement invariante.

INVOLUTIF◆

◆ Lire d'abord les articles « les Ensembles » et « l'Algèbre ».

Soit E un ensemble muni d'une loi de composition interne notée $\divideontimes$ qui admet un élément neutre e ; un élément x de E est involutif si, en composant x avec lui-même, le résultat obtenu est e : $x \divideontimes x = e$.

L'élément x admet donc un symétrique qui est lui-même.

Exemple. Dans l'ensemble $\mathfrak{P}(E)$ des parties d'un ensemble E, muni de la loi de composition interne « différence symétrique » dont l'élément neutre et l'ensemble vide, tout élément est involutif, car, pour toute partie A de E, la différence symétrique de A avec A est vide :

$$A \Delta A = \varnothing.$$

Autre exemple. Les fonctions réelles d'une variable réelle x constituent un ensemble qui est muni d'une loi de composition interne : la loi $\circ$ de composition des applications (dont l'élément neutre est l'application identique : Id). L'application f qui à x réel fait correspondre son opposé $-x$ est involutive, car, pour tout x réel, l'opposé de l'opposé de x est égal à x :

$$-(-x) = x \qquad (\text{donc } f \circ f = \text{Id}).$$

De même, l'application g qui à tout nombre réel x non nul fait correspondre son inverse $\frac{1}{x}$ est aussi involutive, car l'inverse de l'inverse de x est égal à x :

$$\frac{1}{\frac{1}{x}} = x \qquad (\text{donc } g \circ g = \text{Id}).$$

IRRATIONNEL (Nombre) pages 322, 326, 372 et 373, 385.

Un nombre x est un nombre irrationnel s'il n'existe aucun nombre entier (en dehors de 0) dont le produit par x soit un nombre entier.

Par exemple, $\sqrt{2}$ (le seul nombre positif dont le carré soit égal à 2) n'est pas rationnel, car aucun produit $n.\sqrt{2}$ (n entier non nul) n'est entier. Il y a une infinité de nombres irrationnels : $\sqrt{3}$, $\sqrt[3]{5}$, π, e (base des logarithmes népériens) en font partie.

Le développement décimal d'un nombre irrationnel est illimité et non périodique. Prenons l'exemple très connu de $\pi = 3,141\ 59\ \ldots$: développement illimité, pas de périodicité, c'est un nombre irrationnel, et l'exemple de $\frac{19}{28} = 0,678\ 571\ 428\ 571\ 428\ 571\ldots$, qui, lui, n'est pas un nombre irrationnel, mais un nombre rationnel.

ISOCÈLE : voir Polygones

ISOLÉ◆ page 14.

◆ Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles », Intervalle et Topologie.

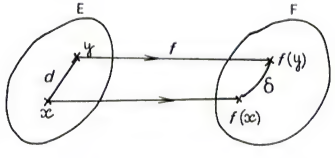
Un point x d'une partie A d'un espace topologique E est isolé s'il existe un voisinage V de x ne contenant aucun point de A autre que x :

$$\mathcal{V}(x) = \{ \text{voisinage de } x \} \quad \exists\, V \in \mathcal{V}(x), \qquad V \cap A = \{ x \}.$$

ISOMÉTRIE, ISOMÉTRIQUE◆

◆ Lire d'abord l'article « les Ensembles » et Distance.

Si E et F sont deux espaces métriques dont les distances respectives sont d et δ, une bijection de E sur F est une isométrie si pour tout couple (x, y) d'éléments de E la distance dans F des images $f(x)$ et $f(y)$ est égale à la distance dans E des éléments x et y : $\forall\, (x, y) \in E^2, \quad \delta[f(x), f(y)] = d(x, y)$.



L'application réciproque f^{-1} de f est également une isométrie. Les deux espaces métriques E et F sont dits isométriques.

ISOMORPHISME : voir Homomorphisme et la page 382.

ITÉRATION (Méthode d')

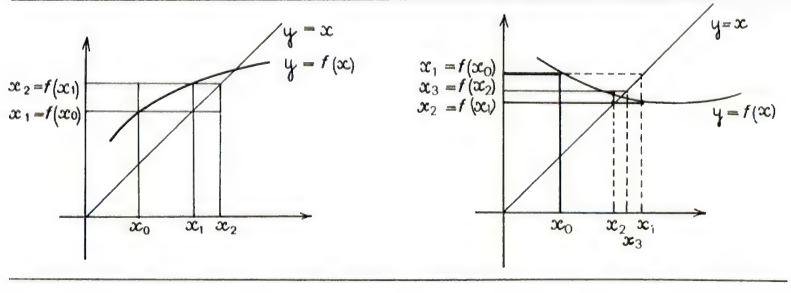
Méthode de calcul permettant de déterminer la solution *unique* d'une équation d'une courbe de la forme $f(x) = x$ (lorsqu'on connaît un nombre k inférieur à 1 tel que la valeur absolue de la dérivée de f soit inférieure à k : $|f'(x)| < k < 1$). Supposons que l'on sache que la solution se trouve dans l'intervalle $[a\ b]$; on prend une valeur x_0 quelconque de cet intervalle et on considère la suite définie par :

$$x_0,\ x_1 = f(x_0), \quad x_2 = f(x_1), \quad x_3 = f(x_2), \quad \ldots$$

On démontre que cette suite est convergente et que sa limite est précisément la valeur x solution de l'équation $f(x) = x$.

Interprétation géométrique

Trouver la solution de l'équation, c'est trouver l'abscisse x du point d'intersection de la droite d'équation $y = x$ avec la courbe d'équation $y = f(x)$.



L

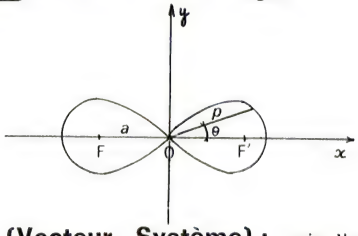
LAPLACIEN : voir Équations aux dérivées partielles.

LEMME

Théorème mis en évidence lors d'une démonstration, mais jugé moins important que d'autres. Il n'en constitue pas moins une étape commode dans le cas d'une démonstration un peu longue.

LEMNISCATE DE BERNOULLI

Lieu des points M du plan dont le produit des distances à deux points fixes F et F', distants de $2a$, est constant et égal à a^2 : $MF \times MF' > a^2$.



L'équation polaire◆ de cette courbe (pôle situé au point double O) est : $\rho^2 = a^2 \cos 2\theta$.

◆ Voir Équation d'une courbe.

Son équation cartésienne◆ est, dans des axes Ox, Oy :

$$(x^2 + y^2)^2 = a^2(x^2 - y^2).$$

◆ Voir Équation d'une courbe.

LIBRE (Vecteur, Système) : voir l'article « l'Algèbre linéaire », pages 67 à 71, 79.

LIÉ (Vecteur, Système) : voir l'article « l'Algèbre linéaire », pages 67 à 71, 79.

LIEU GÉOMÉTRIQUE

Expression employée pour désigner, en géométrie, l'ensemble des points du plan (ou de l'espace) possédant une certaine propriété.

Exemple : le lieu géométrique des points M dont la distance à un point fixe A est égale à R (nombre constant) est le cercle de centre A et de rayon R.

LIMITE : voir l'article « l'Analyse », pages 109, 110, 118 à 124, 149, et les pages 202, 369.

LOGARITHME ◆ pages 159, 303, 327.

◆ Voir aussi Exposant.

Logarithme décimal (ou logarithme de base 10)

Définition « intuitive »

Les nombres 10, 100, 1 000, ..., sont des puissances de 10 :

$$10 = 10^1, \quad 100 = 10^2, \quad 1\,000 = 10^3, \ldots$$

Les logarithmes décimaux de ces nombres sont respectivement les exposants : 1, 2, 3 (ainsi, $\log 10^1 = 1$, $\log 10^2 = 2$, etc.). De même :

$$1 = 10^0, \quad 0{,}1 = \frac{1}{10} = 10^{-1}, \quad 0{,}01 = \frac{1}{100} = 10^{-2}, \ldots$$

Ce sont aussi des puissances de 10 ; leurs logarithmes sont 0, — 1, — 2. Il est donc facile de connaître le logarithme décimal d'une puissance entière de 10, que cette puissance soit positive ou négative. Le logarithme décimal (noté log) de 10^n est n :

$$\boxed{\forall\, n \in \mathbb{Z} : \log 10^n = n}$$

(lire : quel que soit n appartenant à l'ensemble des nombres relatifs (positifs ou négatifs), le logarithme décimal de 10 puissance n est égal à n).
On définit alors de façon analogue le logarithme d'un nombre réel strictement positif x comme étant le nombre réel y tel que x soit égal à 10^y ◆ :

$$\boxed{\log x = y \Longleftrightarrow x = 10^y}$$

◆ Si $y = \log x$ et $y' = \log x'$, cela veut dire que $x = 10^y$ et $x' = 10^{y'}$. Or $x \times x' = 10^y \times 10^{y'} = 10^{y+y'}$, ce qui veut dire que $y + y'$ est le logarithme de $x \times x'$. On a donc la formule : $\log (x \times x') = \log x + \log x'$.

Naturellement, il n'est plus aussi facile de trouver le logarithme d'un nombre qui n'est plus une puissance de 10, comme 57 par exemple. On peut toutefois penser à priori que, puisque 57 est compris entre 10 et 100, son logarithme est compris entre les logarithmes de ces deux nombres, soit 1 et 2 ; de même, si on considère le nombre 0,25, puisqu'il est compris entre 0,1 et 1, son logarithme est compris entre — 1 et 0.
De façon générale, tout nombre réel strictement positif x est compris entre deux puissances successives de 10 (puissances positives ou négatives) : 10^n et 10^{n+1} ; son logarithme est compris entre n et $n + 1$; on peut donc l'écrire $n + m$, où m est une quantité comprise entre 0 et 1. Ainsi :

$\log 57 = 1 + 0{,}755\,87$ ◆ (car $10^1 < 57 < 10^2$)

$\log 0{,}25 = -1 + 0{,}397\,94$ ◆ (car $10^{-1} < 0{,}25 < 10^0$).

◆ Ce nombre est extrait d'une table de logarithme. Nous donnons une telle table à la suite de cet article.

Le nombre entier positif ou négatif n est appelé *caractéristique* du logarithme. Le nombre positif m est appelé *mantisse* du logarithme.
Les tables de logarithmes donnent les mantisses des logarithmes des nombres ayant quatre (ou cinq) chiffres ◆. La caractéristique doit être calculée par l'utilisateur.

◆ 1 + 0,755 87 s'écrit normalement 1,755 87. On convient d'écrire — 1 + 0,397 94 : $\bar{1}$,397 94, en mettant le signe moins au-dessus du 1 pour bien rappeler que, seule, la partie entière est négative, la partie décimale étant positive.

Recherche de la mantisse du logarithme d'un nombre

La propriété suivante est une propriété fondamentale ; elle permet de chercher dans une table la mantisse du logarithme d'un nombre sans s'occuper de la place de la virgule dans ce nombre : *deux nombres formés des mêmes chiffres (par exemple :* 2 457 ; 24,57 ; 0,024 57*) ont des logarithmes dont la mantisse est la même et dont seules les caractéristiques diffèrent.*
Exemple : log 24,57 = 1,390 41.
2 457 est égal au produit de 24,57 par 100 ; le logarithme de 2 457 est donc égal à la somme du logarithme de 24,57 et du logarithme de 100, c'est-à-dire 2 : la partie de log 24,57 située après la virgule n'est pas modifiée par l'addition de 2 ; seule la partie entière l'est :

$\log 2\,457 = 1{,}390\,41 + 2 = 3{,}390\,41.$

◆ Si le nombre dont on cherche le logarithme a plus de quatre (ou cinq) chiffres, il faut faire une interpolation linéaire pour avoir une valeur approchée du logarithme de ce nombre.

0,024 57 est égal à 24,57 divisé par 1 000, c'est-à-dire au produit de 24,57 par 10^{-3}. Le logarithme de 0,024 57 est donc égal à la somme de celui de 24,57 et de celui de 10^{-3}, c'est-à-dire — 3. Seule la partie entière est modifiée par l'addition de — 3 :

$$\log 0{,}024\ 57 = 1{,}390\ 41 - 3 = \bar{2}{,}390\ 41.$$

Recherche de la caractéristique

Pour trouver la caractéristique du logarithme d'un nombre x, seule la place de la virgule dans ce nombre va nous intéresser.

— Si le nombre x est supérieur à 1 et a n chiffres avant la virgule (éventuelle), c'est qu'il est compris entre 10^{n-1} et 10^n (ainsi, un nombre de deux chiffres avant la virgule est compris entre $10 = 10^1$ et $100 = 10^2$). Son logarithme est compris entre $n - 1$ et n ; la caractéristique est donc $n - 1$. Ainsi :

log 525,4 a pour caractéristique 2

log 3 212 434 a pour caractéristique 6, etc.

— Si le nombre est compris entre 0 et 1, appelons n le nombre de zéros figurant dans son écriture (y compris le zéro avant la virgule). Ce nombre est donc compris entre 10^{-n} et 10^{-n+1} (ainsi : 0,004 est plus grand que $0{,}001 = 10^{-3}$, mais plus petit que $0{,}01 = 10^{-2}$). Son logarithme est compris entre $-n$ et $-n + 1$; sa caractéristique est donc $-n$. Ainsi :

log 0,000 2 a pour caractéristique — 4

log 0,25 a pour caractéristique — 1, etc.

Propriétés des logarithmes décimaux

Les propriétés des logarithmes décimaux sont très importantes pour le calcul numérique et se déduisent toutes de la première :

$$\boxed{\log x \times x' = \log x + \log x'}$$

Elles permettent de ramener toutes les opérations de multiplication, division, élévation à une puissance, extraction de racine n-ième à des additions, soustractions, multiplication ou division par un nombre entier :

$$\log \frac{x}{x'} = \log x - \log x'$$

$$\log x^n = n \log x \text{ (avec } n \in \mathbb{Z}\blacklozenge\text{)}$$

$$\log \sqrt[n]{x} = \frac{1}{n} \cdot \log x \text{ (avec } n \in \mathbb{N}\text{)}.$$

◆ Cette formule est encore vraie si n est rationnel.

Logarithme de base quelconque

Au lieu d'utiliser le nombre 10, on aurait pu utiliser n'importe quel nombre a réel positif, différent de 0 et de 1 ◆, pour définir un logarithme de base a dont les propriétés seraient analogues.

Dans les calculs, le logarithme décimal s'est imposé, du fait que notre numération est une numération de base 10.

Un autre logarithme est très utilisé en analyse : c'est le logarithme népérien, ou logarithme de base e ($e = 2{,}718\ 28\ldots$) dont l'abréviation est Log (ou ln). Nous

◆ Différent de 0, car 0^x est toujours égal à 0 pour tout x strictement positif. Différent de 1, car 1^x est toujours égal à 1 pour tout x strictement positif.

allons définir rigoureusement ce logarithme et voir comment, à partir de lui, on peut définir les logarithmes de base a (donc le logarithme décimal en particulier).

Logarithme népérien ◆

◆ Lire d'abord Dérivées primitives ou l'article « l'Analyse », pages 125 et suiv. et Intervalle.

Les primitives F des fonctions $f : f(x) = x^n$ (où n est un nombre entier positif ou négatif) existent et sont données par la formule :

$$F(x) = \frac{x^{n+1}}{n+1} + C$$

(où C est une constante quelconque). Cette formule ne peut pas s'appliquer pour $n = -1$ (le dénominateur est alors nul), c'est-à-dire pour la fonction $f : f(x) = x^{-1} = \frac{1}{x}$.

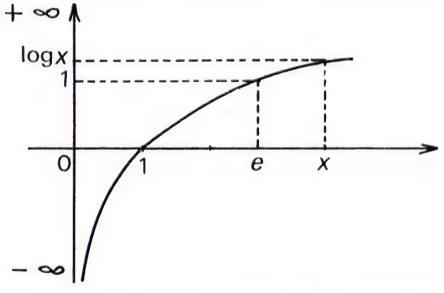
Pourtant, cette fonction est continue sur chacun des intervalles $]-\infty, 0[$ ou $]0, +\infty[$, où $f(x)$ existe et (d'après un théorème concernant les primitives) elle admet donc des primitives sur l'un ou l'autre de ces intervalles. Parmi toutes les primitives existant sur l'intervalle $]0, +\infty[$, il en existe une seule qui prend la valeur 0 si x prend la valeur 1.

Par définition, cette fonction est appelée fonction logarithme népérien. De la définition on déduit immédiatement que Log x n'existe que si x est strictement positif :

$$(\text{Log } x)' = \frac{1}{x} \text{ (dérivée de Log } x)$$

$$\text{Log } 1 = 0.$$

Autres propriétés de cette fonction : c'est une fonction continue ; puisque la dérivée $F'\left(F'(x) = \frac{1}{x}\right)$ est positive, pour tout nombre x positif, le logarithme népérien est une fonction croissante ; on démontre que :



a) la limite de Log x, quand x tend vers $+\infty$, est $+\infty$;
b) la limite de Log x, quand x tend vers 0, est $-\infty$.
La courbe représentative de cette fonction est représentée ci-contre.

Le seul nombre x pour lequel Log x vaut 1 est noté e. Il vaut : 2,718 88...

Si x et x' sont deux nombres réels positifs strictement :

$$\text{Log } x \times x' = \text{Log } x + \text{Log } x'$$

$$\text{Log } \frac{x}{x'} = \text{Log } x - \text{Log } x'.$$

Si n est un entier positif ou négatif, ou même rationnel :

$$\text{Log } x^n = n \text{ Log } x.$$

Si n est un entier positif :

$$\text{Log } \sqrt[n]{x} = \frac{1}{n} \text{Log } x.$$

Si, ayant défini ainsi la fonction Log, on cherche toutes les fonctions f satisfaisant, pour tous nombres réels positifs x et x', à la propriété (que possède la fonction Log) :

$$f(x \times x') = f(x) + f(x'),$$

on trouve que seules les fonctions de la forme $f(x) = k.\text{Log } x$ (où k est une constante) conviennent.
On appelle alors a le nombre réel positif différent de 1 pour lequel $f(x)$ vaut 1 ($f(a) = 1$, or $f(a) = k \times \text{Log } a$, donc $k = \text{Log } a$) et on obtient comme nouvelle expression de f :

$$f(x) = \frac{\text{Log } x}{\text{Log } a}.$$

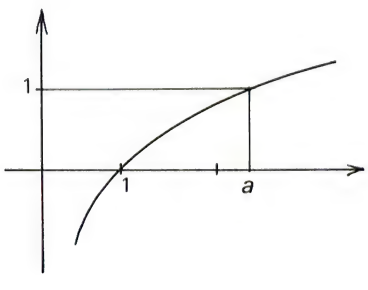
Ce sont ces fonctions f qu'on appelle logarithmes de base a et qu'on note $\log_a$. En particulier, si $a = 10$, on obtient le logarithme décimal $\left(\text{on a donc entre logarithme népérien et logarithme décimal d'un nombre réel positif } x \text{ la relation } \log x = \frac{\text{Log } x}{\text{Log } 10}\right)$.

Les propriétés des logarithmes à base a sont analogues à celles de la fonction Log. La dérivée de $\log_a$ est donnée par :

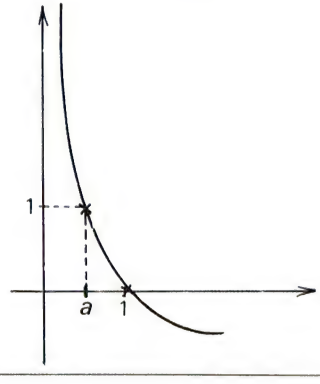
$$(\log_a x)' = \frac{1}{x \text{ Log } a}.$$

Cette dérivée est donc positive si a est plus grand que 1 et négative si a est compris entre 0 et 1 ; les courbes représentatives ont l'allure suivante :

Courbe représentative de la fonction Log a dans le cas où $a > 1$



Courbe représentative de la fonction Log a dans le cas où $0 < a < 1$



MANTISSES DES LOGARITHMES DÉCIMAUX
DES NOMBRES ENTIERS DE 100 à 999 (table à 5 décimales)

N	0	1	2	3	4	5	6	7	8	9
10.	000 00	004 32	008 60	012 84	017 03	021 19	025 31	029 38	033 42	037 43
11.	041 39	045 32	049 22	053 08	056 90	060 70	064 46	068 19	071 88	075 55
12.	079 18	082 79	086 36	089 91	093 42	096 91	100 37	103 80	107 21	110 59
13.	113 94	117 27	120 57	123 85	127 10	130 33	133 54	136 72	139 88	143 01
14.	146 13	149 22	152 29	155 34	158 36	161 37	164 35	167 32	170 26	173 19
15.	176 09	178 98	181 84	184 69	187 52	190 33	193 12	195 90	198 66	201 40
16.	204 12	206 83	209 52	212 19	214 84	217 48	220 11	222 72	225 31	227 89
17.	230 45	233 00	235 53	238 05	240 55	243 04	245 51	247 97	250 42	252 85
18.	255 27	257 68	260 07	262 45	264 82	267 17	269 51	271 84	274 16	276 46
19.	278 75	281 03	283 30	285 56	287 80	290 03	292 26	294 47	296 67	298 85
20.	301 03	303 20	305 35	307 50	309 63	311 75	313 87	315 97	318 06	320 15
21.	322 22	324 28	326 34	328 38	330 41	332 44	334 45	336 46	338 46	340 44
22.	342 42	344 39	346 35	348 30	350 25	352 18	354 11	356 03	357 93	359 84
23.	361 73	363 61	365 49	367 36	369 22	371 07	372 91	374 75	376 58	378 40
24.	380 21	382 02	383 82	385 61	387 39	389 17	390 94	392 70	394 45	396 20
25.	397 94	399 67	401 40	403 12	404 83	406 54	408 24	409 93	411 62	413 30
26.	414 97	416 64	418 30	419 96	421 60	423 25	424 88	426 51	428 13	429 75
27.	431 36	432 97	434 57	436 16	437 75	439 33	440 91	442 48	444 04	445 60
28.	447 16	448 71	450 25	451 79	453 32	454 84	456 37	457 88	459 39	460 90
29.	462 40	463 89	465 38	466 87	468 35	469 82	471 29	472 76	474 22	475 67
30.	477 12	478 57	480 01	481 44	482 87	484 30	485 72	487 14	488 55	489 96
31.	491 36	492 76	494 15	495 54	496 93	498 31	499 69	501 06	502 43	503 79
32.	505 15	506 51	507 86	509 20	510 55	511 88	513 22	514 55	515 87	517 20
33.	518 51	519 83	521 14	522 44	523 75	525 04	526 34	527 63	528 92	530 20
34.	531 48	532 75	534 03	535 29	536 56	537 82	539 08	540 33	541 58	542 83
35.	544 07	545 31	546 54	547 77	549 00	550 23	551 45	552 67	553 88	555 09
36.	556 30	557 51	558 71	559 91	561 10	562 29	563 48	564 67	565 85	567 03
37.	568 20	569 37	570 54	571 71	572 87	574 03	575 19	576 34	577 49	578 64
38.	579 78	580 92	582 06	583 20	584 33	585 46	586 59	587 71	588 83	589 95
39.	591 06	592 18	593 29	594 39	595 50	596 60	597 70	598 79	599 88	600 97
40.	602 06	603 14	604 23	605 31	606 38	607 46	608 53	609 59	610 66	611 72
41.	612 78	613 84	614 90	615 95	617 00	618 05	619 09	620 14	621 18	622 21
42.	623 25	624 28	625 31	626 34	627 37	628 39	629 41	630 43	631 44	632 46
43.	633 47	634 48	635 48	636 49	637 49	638 49	639 49	640 48	641 47	642 46
44.	643 45	644 44	645 42	646 40	647 38	648 36	649 33	650 31	651 28	652 25
45.	653 21	654 18	655 14	656 10	657 06	658 01	658 96	659 92	660 87	661 81
46.	662 76	663 70	664 64	665 58	666 52	667 45	668 39	669 32	670 25	671 17
47.	672 10	673 02	673 94	674 86	675 78	676 69	677 61	678 52	679 43	680 34
48.	681 24	682 15	683 05	683 95	684 85	685 74	686 64	687 53	688 42	689 31
49.	690 20	691 08	691 97	692 85	693 73	694 61	695 48	696 36	697 23	698 10
50.	698 97	699 84	700 70	701 57	702 43	703 29	704 15	705 01	705 86	706 72
51.	707 57	708 42	709 27	710 12	710 96	711 81	712 65	713 49	714 33	715 17
52.	716 00	716 84	717 67	718 50	719 33	720 16	720 99	721 81	722 63	723 46
53.	724 28	725 09	725 91	726 73	727 54	728 35	729 16	729 97	730 78	731 59
54.	732 39	733 20	734 00	734 80	735 60	736 40	737 19	737 99	738 78	739 57

N	0	1	2	3	4	5	6	7	8	9
55.	740 36	741 15	741 94	742 73	743 51	744 29	745 07	745 86	746 63	747 41
56.	748 19	748 96	749 74	750 51	751 28	752 05	752 82	753 58	754 35	755 11
57.	755 87	756 64	757 40	758 15	758 91	759 67	760 42	761 18	761 93	762 68
58.	763 43	764 18	764 92	765 67	766 41	767 16	767 90	768 64	769 38	770 12
59.	770 85	771 59	772 32	773 05	773 79	774 52	775 25	775 97	776 70	777 43
60.	778 15	778 87	779 60	780 32	781 04	781 76	782 47	783 19	783 90	784 62
61.	785 33	786 04	786 75	787 46	788 17	788 88	789 58	790 29	790 99	791 69
62.	792 39	793 09	793 79	794 49	795 18	795 88	796 57	797 27	797 96	798 65
63.	799 34	800 03	800 72	801 40	802 09	802 77	803 46	804 14	804 82	805 50
64.	806 18	806 86	807 54	808 21	808 89	809 56	810 23	810 90	811 58	812 24
65.	812 91	813 58	814 25	814 91	815 58	816 24	816 90	817 57	818 23	818 89
66.	819 54	820 20	820 86	821 51	822 17	822 82	823 47	824 13	824 78	825 43
67.	826 07	826 72	827 37	828 02	828 66	829 30	829 95	830 59	831 23	831 87
68.	832 51	833 15	833 78	834 42	835 06	835 69	836 32	836 96	837 59	838 22
69.	838 85	839 48	840 11	840 73	841 36	841 98	842 61	843 23	843 86	844 48
70.	845 10	845 72	846 34	846 96	847 57	848 19	848 80	849 42	850 03	850 65
71.	851 26	851 87	852 48	853 09	853 70	854 31	854 91	855 52	856 12	856 73
72.	857 33	857 94	858 54	859 14	859 74	860 34	860 94	861 53	862 13	862 73
73.	863 32	863 92	864 51	865 10	865 70	866 29	866 88	867 47	868 06	868 64
74.	869 23	869 82	870 40	870 99	871 57	872 16	872 74	873 32	873 90	874 48
75.	875 06	875 64	876 22	876 79	877 37	877 95	878 52	879 10	879 67	880 24
76.	880 81	881 38	881 95	882 52	883 09	883 66	884 23	884 80	885 36	885 93
77.	886 49	887 05	887 62	888 18	888 74	889 30	889 86	890 42	890 98	891 54
78.	892 09	892 65	893 21	893 76	894 32	894 87	895 42	895 97	896 53	897 08
79.	897 63	898 18	898 73	899 27	899 82	900 37	900 91	901 46	902 00	902 55
80.	903 09	903 63	904 17	904 72	905 26	905 80	906 34	906 87	907 41	907 95
81.	908 49	909 02	909 56	910 09	910 62	911 16	911 69	912 22	912 75	913 28
82.	913 81	914 34	914 87	915 40	915 93	916 45	916 98	917 51	918 03	918 55
83.	919 08	919 60	920 12	920 65	921 17	921 69	922 21	922 73	923 24	923 76
84.	924 28	924 80	925 31	925 83	926 34	926 86	927 37	927 88	928 40	928 91
85.	929 42	929 93	930 44	930 95	931 46	931 97	932 47	932 98	933 49	933 99
86.	934 50	935 00	935 51	936 01	936 51	937 02	937 52	938 02	938 52	939 02
87.	939 52	940 02	940 52	941 01	941 51	942 01	942 50	943 00	943 49	943 99
88.	944 48	944 98	945 47	945 96	946 45	946 94	947 43	947 92	948 41	948 90
89.	949 39	949 88	950 36	950 85	951 34	951 82	952 31	952 79	953 28	953 76
90.	954 24	954 72	955 21	955 69	956 17	956 65	957 13	957 61	958 09	958 56
91.	959 04	959 52	959 99	960 47	960 95	961 42	961 90	962 37	962 84	963 32
92.	963 79	964 26	964 73	965 20	965 67	966 14	966 61	967 08	967 55	968 02
93.	968 48	968 95	969 42	969 88	970 35	970 81	971 28	971 74	972 20	972 67
94.	973 13	973 59	974 05	974 51	974 97	975 43	975 89	976 35	976 81	977 27
95.	977 72	978 18	978 64	979 09	979 55	980 00	980 46	980 91	981 37	981 82
96.	982 27	982 72	983 18	983 63	984 08	984 53	984 98	985 43	985 88	986 32
97.	986 77	987 22	987 67	988 11	988 56	989 00	989 45	989 89	990 34	990 78
98.	991 23	991 67	992 11	992 55	993 00	993 44	993 88	994 32	994 76	995 20
99.	995 64	996 07	996 51	996 95	997 39	997 82	998 26	998 70	999 13	999 57

$\pi \approx 3{,}141\,6$

$\frac{1}{\pi} \approx 0{,}318\,3$

$\lg \pi \approx 0{,}497\,15$

$\lg \frac{1}{\pi} \approx \bar{1}{,}502\,85$

LOGIQUE pages 97, 299, 310, 329, 364, 523.

Situer la logique mathématique

Le lecteur non averti des développements de la logique mathématique moderne, depuis un siècle au plus, pourrait se trouver quelque peu surpris du contenu qui suit. Il n'y verra pas, en effet, la logique présentée comme une dépendance de la philosophie (comme une réflexion sur la mathématique), ni comme une méta-mathématique, fondement naturel, bien que formel, de l'ensemble de l'édifice mathématique.

Sans renier en rien ses origines ni ses caractères purement spécifiques, la logique se veut avant tout une branche de la mathématique moderne, contribuant avec toutes les autres à la solidité et, si possible, à la beauté de l'ensemble. Cet apparent paradoxe mérite explication.

Logique et métamathématique

Pour beaucoup, y compris chez les mathématiciens, la logique reste l'introduction aux mathématiques — ou plutôt, dans la perspective dite moderne, à la mathématique, unifiée justement par ce ciment qu'est la logique. Dans cet esprit, le statut de la logique est généralement celui d'un préambule, d'un préalable, dont souvent on se débarrasse vite mais que parfois on développe : c'est alors le « chapitre zéro ». Plus intuitivement, c'est la règle du jeu, étant implicitement entendu que le jeu ne commence qu'après l'énoncé de la règle. Il y a là pour le moins une grave ambiguïté, qu'il nous faut lever. Les mathématiciens, cela est vrai, pratiquent leur recherche comme un jeu, dont ils doivent évidemment présenter les règles formelles, abstraites. Mais il serait malhonnête de laisser entendre en quelque manière que ces règles sont neutres ou arbitraires. Le mathématicien sait par expérience quelles règles il a intérêt à se fixer s'il veut aboutir à un résultat qui s'intègre sans effort dans l'édifice mathématique. Cela ne signifie absolument pas qu'il n'avait pas le choix de ses conventions, ni qu'un autre choix aurait donné un résultat équivalent. Il y a depuis longtemps — et il y aura encore, souhaitons-le — des mathématiciens qui suivent d'autres règles que celles qu'on présente habituellement comme évidentes : loin d'être inintéressants, leurs résultats sont très enrichissants pour la mathématique.

Ainsi, à la suite de E.J. Brouwer, une école de logiciens, néerlandais surtout, est décidée à n'admettre comme vrai que ce qui est effectivement démontré. Pour eux, il ne suffit pas que la négation d'une propriété soit fausse pour que cette propriété soit déclarée vraie. Nombre de démonstrations classiques deviennent caduques ; certaines peuvent être reprises sous forme positive, d'autres non. Mais tout ce qui est vrai dans cette « logique intuitionniste » est vrai aussi en « logique classique » ◆.

◆ Voir, par exemple, A. Heyting : *les Fondements des mathématiques. Intuitionnisme. Théorie de la démonstration*, coll. « Logique mathématique » (Paris, Gauthier-Villars, 1955).

Une tradition plus ancienne, puisqu'elle remonte à Aristote, reprise sous des formes très diverses, considère que les propositions sont susceptibles d'autres « modalités » (valeurs de vérité) que le « vrai » et le « faux » : « vrai hier, mais faux demain », « absurde », « possible »... La formalisation, relativement récente, s'est effectuée dans deux directions : *a)* dans les *calculs modaux* en logique bivaluée, le langage (classique) est enrichi d'un ou plusieurs symboles représentant les modalités retenues et on raisonne en termes de vrai et de faux ; *b)* dans les *calculs plurivalués*, au lieu de modifier la syntaxe, on peut attribuer aux propositions des valeurs de vérité en nombre supérieur à 2 ◆.

◆ Voir l'article Flou (Mathématique du).

La logique est une mathématique...
et la métamathématique un discours

Mais la logique n'est pas n'importe quelle mathématique. Constituée historiquement à partir des essais successifs faits par les mathématiciens pour fonder les mathématiques, elle a plus ou moins naturellement pris en charge diverses études (les ordres, le calcul booléen ◆, etc.) qui, en toute rigueur, relèveraient plutôt d'autres secteurs de la mathématique moderne. Sans abandonner ces centres d'intérêt, la logique a progressivement tracé sa voie propre, élaboré ses techniques, et s'est constituée en discipline. Cela ne l'a pas coupée des autres spécialités mathématiques, bien au contraire : dans le remodelage permanent des classifications de la mathématique vivante, les frontières de la logique évoluent et évolueront encore ; les échanges se font de plus en plus nombreux entre logique et algèbre, probabilités, topologie, etc. Cette actualité de la logique est bien éloignée de l'image qu'on s'en fait parfois...

◆ Voir Algèbre de Boole.

Logique et théorie des ensembles

L'unification de la quasi-totalité de la mathématique sur les ensembles est un phénomène remarquable. Il ne nous appartient pas de le développer ici. Nous noterons seulement qu'il ajoute de nouvelles ambiguïtés sur le problème, qui est le nôtre, du fondement de la mathématique.

« *Tout est ensemble* »... *une conception à dépasser*

Le fait que les ensembles fassent systématiquement l'objet du chapitre premier des ouvrages modernes pourrait apparaître comme un succédané de cette métamathématique dont nous avons souligné l'aspect problématique. La totalité de l'édifice mathématique pourrait-il être fondé sur cette rubrique? Telle que la question est ici posée, la réponse ne peut être que négative ◆.

Quiconque s'est efforcé de présenter avec rigueur une théorie, quelle qu'elle soit, a constaté que, à moins de faire systématiquement appel à l'intuition — avec tous les risques que cela comporte pour la solidité de la construction —, rien ne peut être fait s'il n'a pas été précisé ce qu'on entend par formule, axiome, théorème, démonstration, etc., et si l'on ne dispose de concepts mathématiques relativement élaborés — usuellement fondés sur... une théorie des ensembles! Cercle vicieux? En un sens : oui. Rien n'est premier en mathématique. Fonder le tout sur la partie semble une entreprise vouée à l'échec. Tout au plus peut-on s'assurer que les divers secteurs de la mathématique sont mutuellement compatibles : les logiciens s'y emploient...

◆ Tout d'abord, l'unification de la mathématique sur les ensembles est d'ores et déjà dépassée. Plusieurs théories en cours de développement, notamment celles qui sont relatives aux catégories, ne peuvent pas se satisfaire du formalisme des ensembles. Ce phénomène ne pourra que s'accentuer. Il faudra trouver à la mathématique un ciment plus fort. Cela ne minimise évidemment en rien la puissance et la beauté de la théorie des ensembles, ni son rôle encore essentiel actuellement. Il importe seulement de ne pas en tirer des conclusions abusives.

Qu'est donc la logique ?

Si la logique n'est rien de ce qu'on voudrait qu'elle soit, qu'est-elle donc? Notre réponse sera double :

— D'une part, la logique est le recueil, toujours actualisé, des tentatives qui ont été faites pour préciser les fondements des mathématiques ; chacune a produit, à défaut d'une explication finale — que nous n'espérons plus trouver ainsi, nous avons dit pourquoi —, une branche neuve et intéressante de travail. Le calcul propositionnel, qu'on trouvera ci-après, en est le meilleur exemple.

— D'autre part, la logique est le développement de recherches théoriques sur ce que peut être une théorie mathématique.

Le calcul propositionnel

Du syllogisme... au calcul booléen

Tout homme est mortel,	*Tout mammifère est vivipare,*	*Aujourd'hui il fait beau,*
Or Socrate est un homme,	*Or le chat est un mammifère,*	*Or il a fait beau hier,*
Donc Socrate est mortel.	*Donc le chat est vivipare.*	*Donc demain il fera beau.*

Ces trois raisonnements sont-ils logiquement valables? Le premier est fort connu depuis la philosophie classique (la scolastique) ; le deuxième, que les logiciens modernes reconnaissent formellement comme acceptable (bien construit), est inintéressant car sa prémisse majeure (première ligne) est fausse (il existe également des mammifères ovipares!) ; le troisième est rejeté à la fois par les logiciens formalistes et... par le bon sens immédiat. Pourquoi des appréciations si différentes sur des raisonnements à première vue très comparables? Dès l'époque classique, les philosophes avaient noté que la validité d'un syllogisme est liée à sa forme et non aux termes qui le composent. Dans notre premier exemple, « Descartes », ou n'importe quel autre nom propre, peut remplacer « Socrate », pourvu que ce soit dans les deux occurrences à la fois ; de même « mortel » peut céder le pas à un autre qualificatif, « homme » à un autre nom de classe d'êtres. Faisons donc abstraction des termes particuliers de ce syllogisme, et notons-en la forme :

Tout (B) *est* (A)
Or (C) *est* (B)
Donc (C) *est* (A).

Dans le formalisme moderne, on écrit :

$$(B) \Longrightarrow (A),\ (C) \Longrightarrow (B) \Longrightarrow (C) \Longrightarrow (A)$$

Lire : B entraîne ($\Longrightarrow$) A, C entraîne B implique ($\Longrightarrow$) C entraîne A.

L'objet du calcul propositionnel est de systématiser ce travail de transcription de tout raisonnement en termes littéraux afin d'en observer la cohérence formelle◆.

◆ En fait, le calcul propositionnel n'y parvient pas ; il faut recourir au calcul des prédicats, qui a lui-même ses limites.

La méthode est simple :
1. On décompose chaque assertion en ses termes les plus élémentaires : les propositions atomiques◆.
2. On convient de représenter par un symbole littéral chaque proposition atomique.
3. On code de manière conventionnelle les termes de liaisons : les connexions.
4. Le résultat est alors traité de manière algébrique.

◆ Seule cette partie du travail nécessite une certaine compréhension du contenu, que nous appellerons « sémantique ».

Ce travail appelle quelques commentaires.

Les variables propositionnelles et les connexions

Les entités sur lesquelles on travaille sont en principe les propositions atomiques. Ce sont des phrases auxquelles peut être attribuée une valeur de vérité, c'est-à-dire la qualification de « vrai » ou de « faux ». La valeur de vérité d'une proposition composée ne doit dépendre que de la valeur des propositions atomiques qui la composent et des connexions qui les lient. Au lieu de considérer que les lettres du calcul propositionnel représentent des propositions, on en est venu à les considérer comme des variables parcourant l'ensemble à deux éléments $\{$ vrai, faux $\}$. On les appelle « variables propositionnelles ». Les connexions sont alors considérées comme des fonctions à une ou plusieurs variables propositionnelles à valeurs dans ce même ensemble à deux éléments $\{$ vrai, faux $\}$ ou $\{$ V, F $\}$. Elles lient donc une, deux ou plusieurs propositions.

Ayant une proposition P connue, quelle connexion◆ pouvons-nous faire agir sur P donnant une nouvelle proposition distincte de P? Il en existe une seule : c'est la *négation* notée ¬P qui à P fait correspondre la proposition ¬P vraie quand P est fausse et fausse quand P est vraie.

◆ Une telle connexion est dite *unaire*.

Ayant deux propositions P et Q connues, quelle connexion◆ pouvons-nous faire agir sur P et Q? Il en existe plusieurs dont voici les principales, traduites par la table de vérité◆ :

◆ Une telle connexion est dite *binaire*.

◆ Pour bien comprendre la suite de l'article, il est conseillé de voir ce mot au dictionnaire.

la *disjonction inclusive* P ou Q, notée $P \vee Q$;

la *conjonction* P et Q, notée $P \wedge Q$;

l'*implication* P entraîne Q, notée $P \Longrightarrow Q$;

l'*équivalence* P équivalent à Q, noté $P \Longleftrightarrow Q$.

Table de vérité

P	¬P
V	F
F	V

P	Q	P ∨ Q	P ∧ Q	P ⟹ Q	P ⟺ Q
V	V	V	V	V	V
V	F	V	F	F	F
F	V	V	F	V	F
F	F	F	F	V	V

Les tables de vérité

Les connexions mathématiques doivent être soigneusement précisées par rapport à celles du langage courant ; celles-ci ont en effet des acceptions extrêmement ambiguës. Ainsi, dans l'expression « les garçons et les filles », la conjonction « et » représente l'adjonction d'une population à une autre ; dans « les garçons grands et forts », la même conjonction « et » représente la restriction d'une population à une sous-population.

Mathématiquement, une *connexion* est considérée comme définie lorsqu'on a précisé sa table de vérité.

Un premier pas est fait dans la construction de propositions composées en appliquant les connexions aux propositions atomiques ; mais on peut continuer la construction en faisant opérer de nouveau ces connexions sur les propositions ainsi obtenues, et ainsi de suite autant qu'on le veut. Par exemple, on peut écrire :

$$P \vee Q, \quad \neg R, \quad Q \wedge S, \quad \text{puis} \quad (P \vee Q) \Longrightarrow (R \wedge S)$$

et, enfin, $\neg R \Longleftrightarrow (P \vee Q) \Longrightarrow (Q \wedge S)$.

La valeur d'une telle proposition complexe est bien déterminée si on connaît les valeurs des propositions atomiques qui la composent : on le vérifie aisément à l'aide des tables des connexions concernées.

Des formules propositionnelles différentes peuvent avoir la même table de vérité.

Donnons quelques exemples :

$X \Longrightarrow Y$, $(\neg X) \vee Y$, $\neg(X \wedge Y)$ sont équivalentes entre elles,

$(X \vee Y) \wedge Z$, $(X \wedge Z) \vee (Y \wedge Z)$ sont équivalentes,

et on pourrait trouver de nombreuses formules encore équivalentes à chacune.

$X \Longrightarrow Y$ est équivalente à $(\neg X) \vee Y$; on dit que la connexion $\Longrightarrow$ se traduit au moyen de $\neg$ et de $\vee$. On montre, par des considérations de combinatoire, que toute connexion se traduit à l'aide de connexions au plus binaires, et même des seules connexions $\neg$, $\vee$, $\wedge$, dont nous avons donné les tables.

La logique et les ordinateurs

Les connexions $\neg$, $\vee$, $\wedge$, considérées en tant qu'opérations sur les formules propositionnelles modulo l'équivalence, possèdent certaines propriétés algébriques élémentaires qui permettent de les comparer à d'autres opérations algébriques connues des mathématiciens. Ainsi, en utilisant le signe $=$ pour noter l'équivalence, on démontre que quels que soient X, Y et Z on obtient des relations telles que: $X \vee Y = Y \vee X$, $\neg(X \vee Y) = \neg X \wedge \neg Y$, etc., toutes identités qui expriment que, comme l'algèbre des parties d'un ensemble, l'algèbre des propositions est une algèbre de Boole ◆. Une autre interprétation, arithmétique celle-ci, s'obtient en remplaçant partout V par le nombre 1, F par 0, et les connexions $\vee$ et $\wedge$ par $+$ et $\times$; cette arithmétique à deux nombres (binaire) ne diffère de l'arithmétique usuelle que par la règle : $1 + 1 = 1$; on s'y habitue facilement. On peut surtout « apprendre » ce calcul aux ordinateurs, dont les circuits élémentaires sont binaires (fermé - ouvert) et se composent comme les connexions.

◆ Cela paraît ici évident, mais devient faux si on s'intéresse à des formules contenant des prédicats. Voir partie « Syntaxe ».

Suggestions pour un « ordinateur » artisanal

circuit « non »

le courant passe en bas lorsque la bobine cesse d'être excitée

circuit « ou »

le courant passe dès qu'un interrupteur est fermé

circuit « et »

le courant ne passe que si les deux interrupteurs sont fermés

Les ordinateurs savent aussi reconnaître si une formule propositionnelle qu'on leur présente est toujours vraie (tautologie), toujours fausse (antilogie) ou si sa valeur dépend des valeurs données aux variables qui y figurent. Il leur suffit évidemment de faire l'essai de toutes les possibilités (2^n calculs pour une formule à n variables).

Syntaxe et sémantique

Nous venons de voir le point de départ de la logique. Le « langage logique » commence à se construire. Nous allons considérer maintenant toutes les possibilités offertes par les règles du jeu de ce langage, la syntaxe d'abord, puis la sémantique.

La syntaxe

Un *énoncé* est une expression composée de symboles convenus : lettres, symboles logiques et signes plus spécifiques et symboles fonctionnels caractéristiques du

langage adopté. La logique précise les règles formatives d'énoncés, et les critères permettent de reconnaître, parmi tous les assemblages de symboles, ceux qui sont admis dans le discours (les formules du langage).
Ces énoncés sont abstraits en ce sens qu'à aucun moment la nature ou même l'existence des termes qui y figurent n'est prise en compte dans cette phase du travail.

Une *théorie* est un ensemble de formules d'un même langage clos par la déduction (c'est-à-dire dont toutes les déductions logiques internes à ce langage ont été tirées).
On appelle *axiomatique* d'une théorie tout ensemble de formules du même langage dont on peut déduire, comme théorèmes, toutes les formules appartenant à la théorie, et elles seules.

Mais certaines formules se déduisent de n'importe quel ensemble d'énoncés (en particulier d'un ensemble vide d'énoncés) ; ce sont les *tautologies* (intuitivement, ce sont les « lapalissades » du langage). Les tautologies forment une théorie (car d'un ensemble de tautologies on ne peut déduire que des tautologies) et cette théorie est incluse dans toute autre théorie.
En revanche, un ensemble de formules est dit « inconsistant » ou « contradictoire », lorsqu'on peut en déduire une formule et sa négation.

La sémantique
La notion syntaxique de langage est, nous l'avons dit, formelle. Mais le jeu qui consiste à construire un langage n'est pas gratuit : il se justifie par le fait qu'il est adéquat à représenter certaines situations qui se rencontrent effectivement en mathématique ; une réalisation d'un langage sera ainsi une situation mathématique que ce langage peut entièrement décrire ; un modèle d'une théorie sera de même une situation dans laquelle tous les énoncés de la théorie sont satisfaits.
Cette notion un peu vague de « situation mathématique » doit être précisée. Il se trouve, et cela est une victoire remarquable de la théorie des ensembles, qu'on peut toujours se ramener à des situations ensemblistes ; précisément, toute théorie consistante admet un modèle dont la base est un ensemble, et toute théorie qui admet un modèle dont la base est un ensemble est consistante. Inutile donc d'aller chercher plus loin. (Rien n'exclut cependant que des logiciens, un jour ou l'autre, trouvent quelque intérêt à le faire...)

C'est donc dans le cadre des ensembles que les notions de réalisation et de modèle vont être définies rigoureusement.

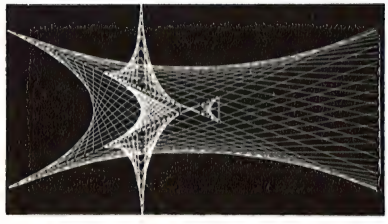
Réaliser un langage, c'est d'abord choisir un ensemble comme base. Les restrictions étant généralement minimes, les réalisations possibles d'un langage forment une famille trop vaste pour être un ensemble : nous l'appellerons classe.
Les réalisations qui rendent vraie une formule particulière sont appelées ses *modèles* ; elles forment une classe qu'on appelle la *classe logique* associée à la formule.
Une formule que toute réalisation rend vraie est dite *universellement valide*.
Doit-on s'étonner de ce qu'on ait pu démontrer que les formules universellement valides sont exactement les tautologies? Cette « coïncidence » (appelée couramment « théorème de complétude ») résulte d'un minutieux travail d'ajustement mutuel entre les règles de la déduction formelle (syntaxique) et celles de l'attribution de valeurs (sémantique). Deux réalisations d'un même langage sont

dites *logiquement équivalentes* (ou *élémentairement équivalentes*) si elles donnent la même valeur à toutes les formules liées du langage. Le corps ordonné des réels et celui des algébriques sont logiquement équivalents ; l'ordre des entiers naturels est logiquement équivalent à l'ordre obtenu en faisant suivre $\mathbb{N}$ d'une chaîne discrète ouverte (du type $\mathbb{Z}$).

Au pays de la syntaxe	Au pays de la sémantique
une tautologie	la classe de toutes les réalisations
une formule (ou une théorie finie axiomatisable)	une classe logique
un ensemble consistant (ou une théorie non contradictoire)	une intersection de classes logiques (réunion de classes d'équivalence logique)
une théorie saturée	une classe d'équivalence logique
une contradiction (la théorie inconsistante)	$\varnothing$

LOI DES GRANDS NOMBRES : voir l'article « les Probabilités », page 437.

LOSANGE : voir Polygones



M

MAJORANT, MAJORÉ : voir l'article « l'Analyse », pages 105 à 107, 113, 117, et la page 379.

MANTISSE : voir Logarithme

MATRICE : voir l'article « l'Algèbre linéaire », pages 72, 81 à 96 et les pages 165, 175, 230, 240, 249, 329, 343, 398, 399, 443, 451, 521, 536, 541, 542.

Matrice conjuguée : voir Conjuguée.

Matrice orthogonale : voir Orthogonale.

Matrice inverse : voir Inverse.

MAXIMAL, MINIMAL

Si E est l'ensemble $\{a, b, c\}$, il y a huit sous-ensembles de E : l'ensemble vide : $\varnothing$; les ensembles à un seul élément : $\{a\}$, $\{b\}$, $\{c\}$; les ensembles à deux éléments : $\{a, b\}$, $\{a, c\}$, $\{b, c\}$; l'ensemble $E = \{a, b, c\}$. L'ensemble de tous ces ensembles est muni d'une relation : l'inclusion, qui est une relation d'ordre.

Considérons l'ensemble A constitué par les sous-ensembles : $\{a\}, \{b\}, \{a,b\}, \{b,c\}$. L'ensemble $\{a\}$ est contenu dans $\{a, b\}$, l'ensemble $\{b\}$ est contenu dans $\{a, b\}$ et dans $\{b, c\}$, mais les ensembles $\{a, b\}$ et $\{b, c\}$ ne sont contenus dans aucun ensemble de A (sauf dans eux-mêmes). On dit que ce sont des éléments maximaux.

De même, $\{a, b\}$ contient $\{a\}$ et contient $\{b\}$; $\{b, c\}$ contient $\{b\}$, mais $\{a\}$ ne contient aucun élément de A (sauf lui-même) et $\{b\}$ ne contient aucun élément de A (sauf lui-même). On dit que $\{a\}$ et $\{b\}$ sont des éléments minimaux.

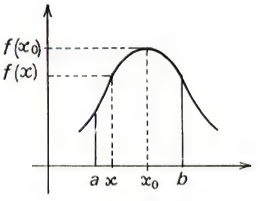
De façon générale, si E est un ensemble muni d'une relation d'ordre notée $\leqslant$, un élément a d'une partie A de E est :

— élément maximal de A s'il n'existe aucun élément x de A qui soit strictement supérieur à a ($a < x$) ;

— élément minimal de A s'il n'existe aucun élément x de A qui soit strictement inférieur à a ($x < a$).

Il peut exister plusieurs éléments maximaux (ou minimaux) dans un ensemble A.

MAXIMUM page 130.

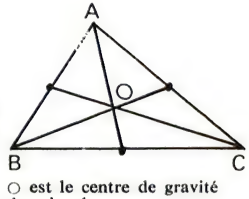


Si f est une fonction réelle d'une variable x, définie sur un intervalle I de $\mathbb{R}$ (ensemble des nombres réels), on dit qu'elle présente un maximum au point x_0 de I s'il existe un intervalle ouvert $]a, b[$ contenant x_0 et tel que, pour tout x de cet intervalle, la valeur de $f(x)$ soit inférieure ou égale à la valeur de $f(x_0)$:

$$f(x) \leqslant f(x_0).$$

Si f est dérivable en x_0, alors la valeur de la dérivée est nulle pour x_0.
Mais le fait que la dérivée s'annule en x_0 ne suffit pas pour qu'il y ait maximum ; il faut que la dérivée s'annule en passant de valeurs positives à des valeurs négatives.

MÉDIANE pages 463, 464.



O est le centre de gravité du triangle

Médianes d'un triangle

Si ABC est un triangle, la médiane issue de A est la droite joignant le point A au milieu du côté BC. On définit de même les médianes issues de B et de C. Les trois droites se coupent en un même point : le centre de gravité du triangle.

En statistiques : voir l'article « Statistiques », page 463.

MÉDIATEUR

Plan médiateur d'un segment (géométrie dans l'espace) : le plan passant par le milieu du segment AB et perpendiculaire à ce segment est appelé plan médiateur du segment AB. C'est aussi l'ensemble des points M de l'espace dont les distances à A et B sont égales.

MÉDIATRICE (en géométrie plane)

D'un segment AB : droite perpendiculaire à la droite AB et qui passe par le milieu du segment AB. Tout point M de la médiatrice de AB est à la même distance de A et de B et, réciproquement, tout point du plan dont les distances à A et B sont égales appartient à la médiatrice de AB.

D'un triangle ABC : les médiatrices des segments AB, BC, CA sont appelées médiatrices du triangle ABC. Elles se coupent en un même point dont les distances à A, B, C sont égales. Il est donc le centre d'un triangle passant par A, par B et par C : c'est le cercle circonscrit au triangle.

MESURE

Nous ne donnerons pas ici de définition exacte de la notion de mesure d'un ensemble, qui est une notion difficile et relativement récente. Nous dirons simplement que la mesure d'un ensemble E, notée m(E), est un nombre réel positif ou nul qui satisfait à certaines conditions, dont les suivantes :

— si E est l'ensemble vide, la mesure de E est nulle ;

— si E n'est pas vide, la mesure de E est positive ;

— si E est contenu dans F, la mesure de E est inférieure ou égale à la mesure de F ;

— si E et F sont deux ensembles disjoints, la mesure de la réunion E U F est égale à la somme des mesures de E et de F.
Les notions de longueur, d'aire, de volume sont des cas particuliers de mesures d'ensembles.

MESURE ALGÉBRIQUE : voir Algébrique et la page 232.

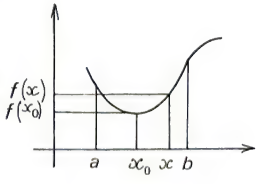
MÉTRIQUE (Espace) : voir Distance et la page 330.

MINIMUM page 130.

Si f est une fonction réelle d'une variable réelle x, définie sur un intervalle I de $\mathbb{R}$ (ensemble des nombres réels), on dit qu'elle présente un minimum au point x_0 de I s'il existe un intervalle ouvert $]a, b[$ contenant x_0 et tel que, pour tout x de cet intervalle, la valeur de $f(x)$ soit supérieure ou égale à la valeur de $f(x_0)$:

$$f(x) \geqslant f(x_0).$$

Si f est dérivable en x_0, alors la valeur de la dérivée est nulle pour x_0. Mais le fait que la dérivée s'annule en x_0 ne suffit pas pour qu'il y ait minimum : il faut qu'elle s'annule en passant de valeurs négatives à des valeurs positives.



MINORANT, MINORÉ : voir l'article « l'Analyse », pages 105 à 107, 113, 117.

MODE : voir l'article « les Statistiques », page 466.

MODULE (Structure de) : voir Vectoriel (Espace) et la page 135.

MODULO : voir Congruence

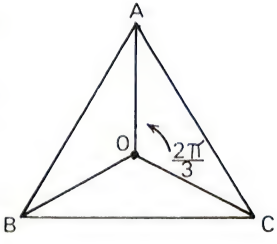
MONOGÈNE

Soit ABC un triangle équilatéral dont on appelle O le centre de gravité et soient r_1, r_2, r_3 les trois rotations de centre O suivantes :

— r_1 : rotation d'angle nul, dans laquelle A se transforme en A, B en B, C en C ;

— r_2 : rotation d'angle $\frac{2\pi}{3}$, dans laquelle A se transforme en B, B en C, C en A ;

— r_3 : rotation d'angle $\frac{4\pi}{3}$, dans laquelle A se transforme en C, B en A, C en B.



On peut composer ces rotations par la loi ∘ de composition des rotations, et on constate que le composé de deux quelconques de ces trois rotations est encore une d'entre elles (par exemple, si on compose r_3 et r_2, c'est-à-dire si on fait successivement r_2 puis r_3, on trouve r_1, car A se transforme d'abord en B par r_2, lequel se transforme en A par r_3 ; B se transforme d'abord en C par r_2, lequel se transforme en B par r_3 ; C se transforme en A par r_2, lequel se transforme en C par r_3 ; partis de A, B, C dans cet ordre, nous trouvons à l'arrivée A, B, C dans cet ordre, c'est-à-dire qu'on a effectué r_1).
Cette loi de composition est donc une loi de composition interne dans l'ensemble R

de ces trois rotations ◆. On peut démontrer que l'ensemble R muni de cette loi est un groupe commutatif. On constate alors que les trois éléments de R peuvent être obtenus à partir d'un seul d'entre eux ; en effet, chaque élément de R peut être obtenu à partir de r_2 uniquement :

— la rotation r_3 peut être obtenue en faisant deux fois successives la rotation r_2, donc :

$$r_3 = r_2 \circ r_2 \text{ ou } (r_2)^2 \text{ (« } r_2 \text{ au carré »)};$$

— la rotation r_1 peut être obtenue en faisant trois fois successives la rotation r_2, donc :

$$r_1 = r_2 \circ r_2 \circ r_2 \text{ ou } (r_2)^3 \text{ (« } r_2 \text{ au cube »)}.$$

De la même façon, chaque élément de R peut être obtenu à partir de r_3 uniquement, car :

$$r_2 = r_3 \circ r_3 = (r_3)^2$$

$$r_1 = r_3 \circ r_3 \circ r_3 = (r_3)^3.$$

Un tel groupe, dont chaque élément est une puissance d'un élément a du groupe, est appelé groupe monogène (on dit aussi qu'il est engendré par a).

◆ La table de Pythagore de la loi ○ dans R est la suivante :

○	r_1	r_2	r_3
r_1	r_1	r_2	r_3
r_2	r_2	r_3	r_1
r_3	r_3	r_1	r_2

MONOÏDE : voir l'article « l'Algèbre », pages 27 et 28.

MONÔME pages 206, 408.

Un monôme est un produit d'éléments de $\mathbb{R}$ (ensemble des nombres réels). Par exemple : ab, axy, $-a^2bxy$, etc. sont des monômes.
Le produit de deux monômes est un monôme, alors que la somme de deux monômes n'est pas, en général, un monôme.
Cette définition est valable pour n'importe quel ensemble ayant une structure d'anneau◆ commutatif.

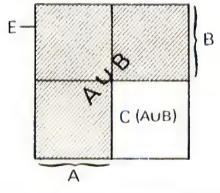
◆ Voir l'article « l'Algèbre ».

MONOTONE : voir l'article « l'Analyse », pages 112 et 113, 122 et 124, 130.

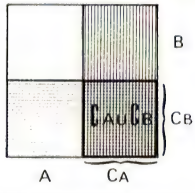
MORGAN (Lois de) pages 271, 329.

Si A et B sont deux parties d'un même ensemble E, le complémentaire de leur réunion A ∪ B est égal à l'intersection du complémentaire de A et du complémentaire de B : $\complement(A \cup B) = \complement A \cap \complement B$.

Complémentaire de AUB



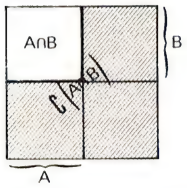
Intersection du complémentaire de A et du complémentaire de B



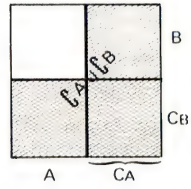
De la même façon, le complémentaire de l'intersection des deux parties A et B est égal à la réunion du complémentaire de A et du complémentaire de B :

$$\complement(A \cap B) = \complement A \cup \complement B.$$

Complémentaire de A∩B



Réunion du complémentaire de A et du complémentaire de B



Ces deux propriétés, liant les opérations (entre parties d'un même ensemble E) de complémentaire, réunion et intersection, sont appelées lois de Morgan.
Il existe, pour les opérations (liant des propositions logiques) de négation, de disjonction (« ou ») et de conjonction (« et »), des relations analogues qui sont aussi appelées lois de Morgan :

— la négation de la proposition « p ou q » est équivalente à la proposition « négation de p et négation de q » :

non (p ou q) $\Longleftrightarrow$ (non p) et (non q) ;

— la négation de la proposition « p et q » est équivalente à la proposition « négation de p ou négation de q » :

non (p et q) $\Longleftrightarrow$ (non p) ou (non q).

MORPHISME : voir Homomorphisme et les pages 40 à 45, 49 à 51.

MOYENNE : voir l'article « les Statistiques », pages 464, 508, 510.

MULTIPLICATION, PRODUIT◆, FACTEUR pages 22, 25, 202, 368 à 389.

◆ Lire aussi l'article « les Ensembles », page 270.

Si A et B sont deux ensembles dont les cardinaux respectifs sont m et n, on appelle produit de ces deux entiers naturels et on note $m.n$ le cardinal du produit cartésien $A \times B$ (m, n et $m.n$ sont trois entiers naturels) ; la loi de composition interne dans $\mathbb{N}$ (entiers naturels) ainsi définie est appelée multiplication. Cette loi est associative, commutative, telle que 1 soit élément neutre, et cette loi est distributive par rapport à l'addition ◆.

◆ Voir l'article « l'Algèbre », pages 21 à 27.

De façon générale, on a l'habitude de noter . et d'appeler multiplication toute loi de composition interne dans un ensemble E qui est associative et d'appeler le composé de deux éléments x et y de E produit de ces deux éléments.
On peut définir le composé de n éléments $x_1 \,.\, x_2 \,.\, \ldots \,.\, x_n$, puisque la loi est associative, et on l'appelle produit de ces n éléments ; chaque élément est appelé un *facteur* du produit.

NO

NATUREL (Nombre) : voir Entier (nombre) et les pages 17, 24, 29, 244, 256 et 257, 264, 366 à 369, 378 et 379.

NÉGATIF page 253.

Nombre négatif : nombre inférieur à 0◆.

◆ Lire également Positif en remplaçant « supérieur » par « inférieur ».

NÉPÉRIEN : voir Logarithme

NÉPHROÏDE : voir Cycloïde

NON NÉGATION, page 271.

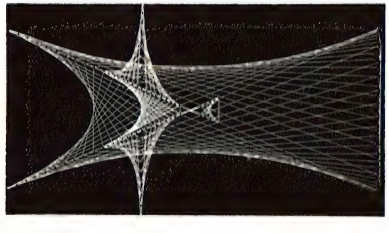
Si p est une proposition, la négation de la proposition p est la proposition qui est vraie quand p est fausse et fausse quand p est vraie. On note cette proposition :

non p ou $\neg p$.

Sa table de vérité est représentée ci-contre.

p	non p
V	F
F	V

NOMBRES : voir l'article pages suivantes et les pages 184, 262 à 270, 324, 326.



Les nombres

par André Warusfel

Les mathématiques ont commencé par deux études intuitives portant sur deux concepts fondamentaux : le *point*, ouvrant sur la géométrie, la structuration de l'espace, prélude indispensable à toute connaissance de la matière et à toute physique, et le *nombre*. Sans nombre, pas de mesure. Ce qui ne signifie pas que sans lui on ne puisse pas raisonner logiquement. Nous savons aujourd'hui construire de vastes théories dans lesquelles n'apparaît aucune mesure : la topologie générale, par exemple, peut se passer de toute considération numérique pendant de longs chapitres. Cela infirme la vieille définition des mathématiques, « étude des grandeurs mesurables ». Cela dit, le nombre garde évidemment toute son importance dans les mathématiques contemporaines, mais son utilisation n'est plus tout à fait la même ; il est de plus en plus généralisé, placé dans des ensembles de plus en plus vastes.

Dans un premier stade, nous nous contenterons d'explorer, rapidement, les nombres dans un deuxième, utilisant la théorie des ensembles, nous étudierons de manière plus détaillée les généralisations successives qui, de l'ensemble des entiers, nous amèneront à l'ensemble des complexes.

André Warusfel.
Né le 1er décembre 1936. Ancien élève de l'École normale supérieure; professeur de mathématiques spéciales au lycée Louis-le-Grand ; il a publié un *Dictionnaire raisonné de mathématiques*, un livre consacré aux structures algébriques finies ainsi que des ouvrages d'enseignement et de vulgarisation des mathématiques.

Le nombre entier

Le nombre le plus simple, le seul à être connu de la plupart de nos contemporains, est évidemment le nombre entier *strictement positif* (1, 2, 3, 4, ..., *n*, ...), élément de l'ensemble que l'on note $\mathbb{N}^*$. $\mathbb{N}^* = (1, 2, 3, 4, \ldots, n, \ldots)$◆ (les entiers naturels).

◆ $\mathbb{N}$ = ensemble des entiers naturels, y compris 0. $\mathbb{N}^*$ = ensemble des entiers naturels, sauf 0.

Le nombre entier apparaît dans l'opération de comptage et d'addition

L'exemple célèbre du berger comptant ses moutons montre de quoi il s'agit ; l'homme récite une litanie qu'il connaît par cœur, fran-

chissant une étape nouvelle chaque fois qu'il remarque un nouvel animal. Quand il a épuisé son troupeau, le dernier mot prononcé correspond au nombre de ses bêtes.
Le premier problème consiste donc à créer cette litanie. Ses premiers éléments sont arbitraires. Dans notre langue, aucun des mots de « un » à « seize » n'est relié à un autre. Ensuite viennent « dix-sept » et les autres. Quelques termes encore pour nommer les dizaines — dont le curieux « quatre-vingts » — « cent », « mille », « million », « milliard », suffisent alors pour qu'en les combinant on puisse largement nommer tous les nombres usuels.
Le caractère très ancien de ce langage particulier est attesté par les bizarreries qui ne nous frappent guère, mais qui sont surprenantes : pourquoi dit-on « vingt et un » et « vingt-deux »? Les étrangers ont aussi leurs particularités propres.
Pour la plupart des hommes, cela suffit grandement. Certaines peuplades ne connaissent même qu'une arithmétique sommaire, où il n'y a qu'un nombre (*sic*) fini de nombres : par exemple, 1, 2, 3, et *b* (*b* pour « beaucoup »). Toute quantité supérieure à 3 est qualifiée de « beaucoup ». Voici ci-contre les tables d'addition et de multiplication.

+	1	2	3	*b*
1	2	3	*b*	*b*
2	3	*b*	*b*	*b*
3	*b*	*b*	*b*	*b*
b	*b*	*b*	*b*	*b*

×	1	2	3	*b*
1	1	2	3	*b*
2	2	*b*	*b*	*b*
3	3	*b*	*b*	*b*
b	*b*	*b*	*b*	*b*

Naturellement, le scientifique ne saurait se contenter d'un matériel relativement pauvre comme $\mathbb{N}^*$. Même si les problèmes, étudiés en théorie des nombres, posés par cet ensemble sont redoutables, ils ne peuvent être réellement compris et résolus qu'avec un arsenal mathématique très complexe, englobant d'autres nombres que les entiers eux-mêmes.
Par exemple, il est connu depuis longtemps (Fermat, au XVII^{e} siècle) que le seul entier x tel que $(x^3 - 2)$ soit un carré parfait est égal à 3 ◆ ; pour démontrer qu'il n'y en a pas d'autres, il faut de solides connaissances ou du génie! Il existe de très nombreuses questions de ce genre, à l'énoncé très élémentaire et à la résolution des plus délicates, et encore bien davantage de questions sans réponse, que l'on appelle des *problèmes ouverts*. Citons-en un seul, à l'aspect bénin : existe-t-il un triplet d'entiers (x, y, z) tel que $x^3 + y^3 - z^3 = 30$? Laissant de côté les redoutables énigmes de la théorie des nombres, nous pouvons dire que, dans la réalité concrète, le maniement des nombres entiers est très simple et commode. Pratiquement, on ne travaille que sur eux ; par construction même, les ordinateurs ne connaissent que les nombres entiers ; écrire un nombre non entier comme $\frac{2}{5}$ sous la forme 0,4 permet, par une astuce technique, de travailler sur lui comme sur l'entier 4, quitte ensuite à ramener la virgule à sa place. C'est ainsi que, pour calculer un produit tel que :

$$3{,}25 \times 0{,}0164,$$

◆ En effet, $3^3 - 2 = 27 - 2 = 25 = 5^2$

on calcule en fait le produit :

$$\begin{array}{r} 325 \\ \times\ 164 \\ \hline 1300 \\ 1950 \\ 325 \\ \hline 53300 \end{array}$$

que l'on ramène ensuite à 0,0533.
Il est d'ailleurs intéressant de remarquer que ces techniques de multiplication, addition, division et soustraction sont bien loin d'être immédiates. Deux faits le prouvent : d'une part, la date tardive de leur mise au point (au Moyen Age, être capable d'effectuer une division était un signe de compétence certain) ; d'autre part, le fait que neuf personnes sur dix, au moins, qui utilisent journellement ces techniques, sont totalement incapables de les justifier, car il s'agit pour elles d'automatismes.

Finalement, nous voyons que l'univers $\mathbb{N}$ des nombres entiers, que l'on fait commencer aujourd'hui par 0 ($\mathbb{N} = \{0, 1, 2, 3, \ldots\}$), muni de ses lois usuelles ($+$, $\times$), est à la fois extrêmement familier à chacun et très complexe, qu'il s'agisse de ses propriétés cachées qu'explore la théorie des nombres ou de sa définition même. C'est vraiment la base des mathématiques ; le langage « Lincos » (de « lingua cosmica »), mis au point pour converser avec d'éventuels extra-terrestres, ne commence-t-il pas par des énumérations d'entiers pour essayer de trouver le contact ? Pour enrichir $\mathbb{N}^*$, on peut procéder, comme l'histoire nous le montre, dans deux directions différentes.

Le nombre fractionnaire

La première consiste à inventer de nouveaux nombres destinés à rendre possibles les divisions par des entiers. C'est ainsi que sont nées les fractions les plus usuelles telles que $\frac{1}{2}$, $\frac{1}{3}$ ou $\frac{3}{3}$. Une illustration familière en est donnée par les fameux partages de tartes. Cet aspect intuitif ne doit pas faire illusion ; la construction complète de l'ensemble des fractions à éléments entiers — que l'on note $\mathbb{Q}_+^*$ en mathématiques *(nombres rationnels strictement positifs* ◆*)* — est assez délicate. Des égalités évidentes comme $\frac{3}{2} = \frac{6}{4}$ y jouent un rôle essentiel. L'ensemble obtenu est sans doute le plus ancien groupe ◆ considéré par les mathématiciens. Ce groupe est un groupe multiplicatif. Il contient la suite remarquable : $\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots,$

◆ $\mathbb{Q}$ = ensemble des nombres rationnels.
$\mathbb{Q}^*$ = ensemble des nombres rationnels, 0 exclu.
$\mathbb{Q}_+^*$ = ensemble des nombres rationnels strictement supérieurs à 0.
$\mathbb{Q}_-^*$ = ensemble des nombres rationnels strictement inférieurs à 0.

◆ Voir l'article « l'Algèbre », page 28.

qui a constitué le premier exemple de limite et de continuité, concepts qui ont connu une fortune éclatante.

Le zéro

L'autre extension naturelle conduit, elle aussi, à un groupe important. Elle passe d'abord par une invention fondamentale, celle du zéro. Elle est peu naturelle, car, si le nombre « trois » peut être conçu aisément à l'aide d'ensembles comme $\{a, b, c\}$ et tous les ensembles qui peuvent être mis en bijection avec lui, la notion de zéro repose sur celle d'ensemble vide, $\varnothing$ ou $\{\ \}$, ce qui représente une abstraction délicate ◆. Le zéro et les nombres négatifs comme $-1, -2, -3$, etc. répondent à un besoin arithmétique. De même que l'on a inventé $\frac{2}{3}$ pour résoudre l'équation $3x = 2$, un nombre comme -1 a été introduit pour donner une solution à l'équation $3 + x = 2$.

◆ Lire à ce sujet l'article « les Ensembles », page 263.

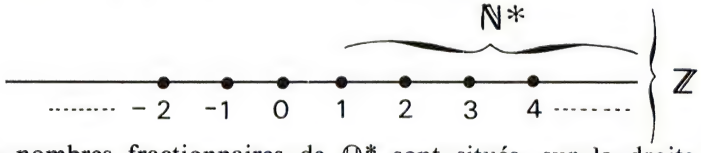
Le nombre entier relatif

Cette construction de l'ensemble $\mathbb{Z} = \{\ldots, -n, \ldots, -1, 0, 1, 2, 3, \ldots, n, \ldots\}$ (ensemble des entiers relatifs ◆) est bien connue maintenant des jeunes enfants, car elle est enseignée à partir de la classe de cinquième. Elle conduit à des nombres dont l'utilité est certaine ; il suffit de penser à leur utilisation pour mesurer une température ou une altitude.

◆ Certains disent « algébriques » ou « rationnels », mais ces termes prêtent à ambiguïté.

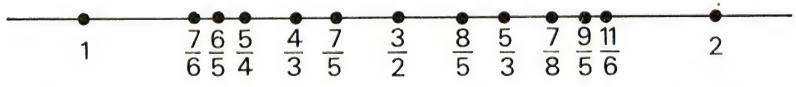
Un nombre entier relatif permet la soustraction

Mais pendant longtemps ces nombres ont été considérés avec méfiance ; c'est ainsi qu'à la Renaissance, époque de virtuosité dans le calcul algébrique, on distinguait, dans la résolution de l'équation $x^3 + px + q = 0$, différents cas suivant les *signes* des nombres p et q. Une égalité comme $a + b = a - ((a - a) - b)$, dont se joue avec facilité n'importe quel écolier, correspond en réalité à une connaissance profonde du groupe additif qu'est $\mathbb{Z}$, ou tout autre groupe additif. $\mathbb{Z}$ est beaucoup plus facile à visualiser que le groupe $[\mathbb{Q}_+^*$ des fractions $\frac{m}{n}$, où m et n appartiennent à $\mathbb{N}^*]$; l'image la plus simple de $\mathbb{Z}$ est évidemment la suivante :



Les nombres fractionnaires de $\mathbb{Q}_+^*$ sont situés, sur la droite qui nous a servi de support, de manière si complexe qu'on ne saurait les

représenter ; les points nécessaires seraient en nombre infini sur n'importe quelle partie de la droite du côté des entiers positifs ; entre deux points quelconques en figure au moins un troisième, et même une infinité : on traduit cela en disant que les nombres fractionnaires forment un ensemble « dense ». Le schéma ci-dessous :



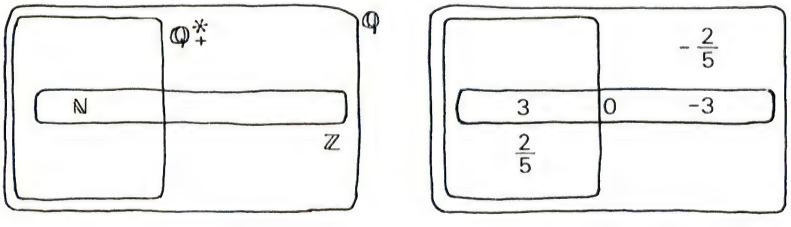
donne une idée, bien médiocre, du fourmillement des rationnels de dénominateur inférieur à 6 compris entre 1 et 2. Les deux extensions ci-dessus, fondées chacune sur le désir de rendre possibles des opérations telles que la soustraction ou la division, souvent impossibles dans $\mathbb{N}^*$ seul, sont indépendantes l'une de l'autre bien que fondées sur des schémas abstraits voisins. Les deux groupes obtenus, $\mathbb{Q}_+^*$ et $\mathbb{Z}$, n'ont en commun que les éléments $\mathbb{N}^*$.

Le nombre rationnel

Il était naturel de rechercher une troisième extension qui prolonge à la fois $\mathbb{Q}_+^*$ et $\mathbb{Z}$. Cette extension est notée $\mathbb{Q}$; ses éléments sont les nombres rationnels, obtenus en ajoutant à $\mathbb{Q}_+^*$ le nombre 0 et l'ensemble des opposés des éléments de $\mathbb{Q}_+^*$ (l'ensemble de ces opposés étant noté $\mathbb{Q}_-^*$) ◆. On obtient les schémas suivants, où figure un élément caractéristique de chacun des quatre ensembles : $\mathbb{N}^*$, $\mathbb{Q}_+^*$, $\mathbb{Z}$ et $\mathbb{Q}$.

◆ Exactement comme on est passé de $\mathbb{N}^*$ à $\mathbb{Z}$ en ajoutant le zéro et les opposés des éléments de $\mathbb{N}^*$, c'est-à-dire les entiers négatifs.

Le nombre rationnel permet la division



$\mathbb{Q}$ est, comme $\mathbb{Z}$, un groupe additif. Non seulement l'addition y possède toutes les propriétés usuelles (par exemple, $a + b = b + a$), mais la soustraction y est toujours possible (par exemple, $\left(-\frac{2}{3}\right) = x + \left(\frac{1}{4}\right)$ a pour solution $x = -\frac{11}{12}$). Si l'on enlève 0 à $\mathbb{Q}$, ce qui donne l'ensemble $\mathbb{Q}^*$ des rationnels non nuls, on obtient un groupe multiplicatif, comme $\mathbb{Q}_+^*$. Non seulement la multiplication

y possède toutes les propriétés usuelles (par exemple, $ab = ba$), mais la division y est toujours possible (par exemple, $\left(-\frac{2}{3}\right) = x \times \left(\frac{1}{4}\right)$ a pour solution $x = -\frac{8}{3}$).

De plus, l'addition et la multiplication restent, comme dans $\mathbb{N}^*$, liées par une relation fondamentale :

$$a(b + c) = ab + ac.$$

Les trois extensions considérées, engendrant $\mathbb{Q}_+^*$, $\mathbb{Z}$ et $\mathbb{Q}$ à partir de $\mathbb{N}^*$, qui est la base de tous les ensembles de nombres, respectent donc les lois, découvertes empiriquement, de l'addition et de la multiplication de $\mathbb{N}^*$. Mais le calcul, dans $\mathbb{Q}$, est beaucoup plus simple que dans $\mathbb{N}^*$, car les procédés utilisés ont permis de lever presque toutes les impossibilités qui gouvernent $\mathbb{N}^*$, où l'on ne peut :
— diviser a par b que si a est un multiple de b ;
— et soustraire c de d que si d est supérieur à a.
Le seul interdit qui subsiste concerne la division : on ne peut diviser par 0.

Comme pour $\mathbb{Q}_+^*$, $\mathbb{Q}$ n'est pas représentable sur une droite. Entre les points qui figurent les éléments de $\mathbb{Z}$, il faut imaginer une infinité de points, de chaque côté de 0, dans un entassement considérable ; $\mathbb{Q}$ est dense comme $\mathbb{Q}_+^*$, dont il n'est, en quelque sorte, que le doublé, 0 mis à part. Si cet adjectif n'avait, en mathématiques, un autre sens très précis, nous dirions que $\mathbb{Q}$ est extrêmement compact sur la droite.

Le nombre décimal

Il ne faut pas croire que $\mathbb{Q}$ soit le seul à avoir cette propriété ; il en est de même de $\mathbb{D}$, qui est l'ensemble des nombres décimaux (ceux qui peuvent s'écrire sous une forme du genre — 5,132 4 avec un nombre fini de décimales après la virgule).

Tout décimal est rationnel ; la réciproque est fausse

Ainsi $\frac{1}{3} = 0{,}333\,3 \ldots 333 \ldots$ n'est pas décimal mais rationnel.

$\mathbb{D}$ est moins riche que $\mathbb{Q}$; le quotient de deux décimaux n'est que rarement un décimal. *Exemple :* le quotient de 0,4 par 0,6 n'est pas un nombre décimal.

$$0{,}4 : 0{,}6 = 0{,}666\,666 \ldots$$

(nombre infini de décimales après la virgule).

Alors que $\mathbb{Q}$ a la structure de corps ◆ (ensemble ou les quatre opérations, sauf la division par 0, sont possibles), $\mathbb{D}$ n'est qu'un anneau ◆ (ensemble où l'addition, la soustraction et la multiplication sont

◆ Lire l'article « l'Algèbre », page 53.

◆ Lire l'article « l'Algèbre » page 46.

possibles), comme $\mathbb{Z}$. Pourtant, à la différence de $\mathbb{Z}$ et à la ressemblance de $\mathbb{Q}$, $\mathbb{D}$ est un ensemble dense. Entre deux rationnels, et à fortiori entre deux décimaux, il existe une infinité de décimaux.

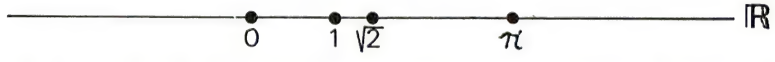
Le nombre irrationnel

La richesse de $\mathbb{Q}$, sa densité, la simplicité des calculs qui y sont pratiqués pourraient faire penser que l'ensemble est achevé. Les Grecs, qui connaissaient fort bien au moins $\mathbb{Q}_+^*$, le pensaient ; ils durent s'apercevoir assez vite qu'il n'en était rien et que $\mathbb{Q}$ présentait, en réalité, de nombreuses lacunes.
Ils pensaient ainsi que $\mathbb{Q}$ contenait un nombre, approximativement égal à $1{,}414 = \frac{707}{500}$ dans notre système de numération, qui, élevé au carré, satisfait à l'équation $x^2 = 2$. En réalité, $(1{,}414)^2 = 1{,}999\,396$. Toute tentative de mettre ce nombre sous la forme $x = \frac{m}{n}$, parfaitement possible lorsqu'il ne s'agit que de trouver une bonne approximation — ainsi $\left(\frac{99}{70}\right)^2 = 2{,}000\,204\ \ldots$ —, est irréalisable si l'on désire un résultat exact. La démonstration de ce fait est due aux Grecs eux-mêmes ; elle ruina la prétention de ramener tous les nombres ◆ à des nombres rationnels. Une nouvelle extension de $\mathbb{Q}$ est alors indispensable. D'autres nombres importants, comme π, n'appartiennent d'ailleurs pas à $\mathbb{Q}$; bien que l'on ait été incapable de le prouver dans l'Antiquité, ce résultat était pressenti et rendait également urgente la complétion de l'ensemble $\mathbb{Q}$.

◆ Y compris ceux dont la géométrie a besoin, comme $\sqrt{2}$, qui est le rapport de la diagonale d'un carré à son côté, d'après le théorème de Pythagore.

Le nombre réel

Cette nouvelle extension de $\mathbb{Q}$ est appelée l'ensemble des nombres réels ; il est noté $\mathbb{R}$. Comme les entiers, les réels sont susceptibles d'une représentation intuitive très simple ; les points associés aux nombres réels recouvrent toute la droite :



On a représenté ici quelques-uns parmi les nombres réels les plus importants ; d'abord 0 et 1 qui, comme tous les rationnels, sont des nombres réels ; la position des points marqués 0 et 1 détermine totalement celle des autres points. Ensuite, $\sqrt{2}$ et π, qui sont les nombres irrationnels (c'est-à-dire réels et non rationnels) les plus marquants.
Il ne faudrait pas croire qu'il n'existe que quelques irrationnels, créés pour les besoins de la cause pour résoudre quelques équations

comme $x^2 = 2$. En réalité, les nombres irrationnels forment, comme $\mathbb{Q}$, un ensemble dense.

Entre deux nombres réels, il existe une infinité de rationnels et une infinité d'irrationnels

Bien plus, il existe infiniment plus d'irrationnels que de rationnels ; ou encore presque tous les nombres sont irrationnels, bien que les rationnels soient évidemment les plus utilisés dans la pratique. La densité d'un ensemble comme $\mathbb{Q}$, ou comme $\mathbb{D}$, fait que l'on peut toujours remplacer un nombre réel par un rationnel ou même un décimal, avec une erreur aussi petite que l'on veut.
Parmi toutes les extensions envisagées ici, c'est l'étape de $\mathbb{Q}$ à $\mathbb{R}$ qui est cruciale. Nous ne pouvons dire ici que quelques mots à ce sujet. Pour définir $\mathbb{R}$, il suffit de connaître tous les réels entre 0 et 1. Un tel nombre sera simplement défini comme une écriture :

$$0, d_1 d_2 \ldots d_n \ldots,$$

où les entiers d_i appartiennent à l'ensemble $\{0, 1, 2, \ldots, 8, 9\}$. Il faut toutefois prendre quelques précautions. Ainsi est-on obligé, pour des raisons techniques, de considérer comme égaux des nombres tels que :

$$0{,}237\,000 \ldots 0 \ldots \text{ et } 0{,}236\,999\,9 \ldots 9 \ldots$$

Les nombres réels forment, comme les nombres rationnels, un corps ◆. Les quatre opérations y sont possibles et obéissent aux mêmes axiomes que dans $\mathbb{N}^*$. Il existe dans $\mathbb{R}$ de nombreux groupes, comme $\mathbb{R}$ lui-même, comme groupe additif, ou $\mathbb{R}^*$ — c'est-à-dire les réels non nuls —, comme groupe multiplicatif. Il en existe bien d'autres, comme le groupe additif des nombres de la forme $(a + b\sqrt{2})$, où a et b sont entiers, le groupe multiplicatif $\mathbb{R}_+^*$ des réels strictement positifs, etc.

◆ Voir l'article « l'Algèbre », page 53.

Le nombre algébrique ; le nombre transcendant

Nous devons quand même signaler, très brièvement, un corps important, intermédiaire entre $\mathbb{Q}$ et $\mathbb{R}$, que l'on peut noter $\mathbb{A}$ (corps des nombres algébriques). Il s'agit des réels x tels qu'il existe des entiers relatifs $n, a, b, c, \ldots, l$ tels que :

$$0 = ax^n + bx^{n-1} + cx^{n-2} + \ldots + l$$

Tout rationnel est algébrique. Des irrationnels comme $\sqrt{2}$ sont algébriques, car $0 = (\sqrt{2})^2 + (-2)$. D'autres irrationnels ne sont pas algébriques ; on les dit transcendants. C'est le cas de π et d'un autre nombre important en analyse, le nombre e ◆ (approximativement égal à 2,718 28). A vrai dire, presque tous les nombres sont transcendants. Tous les ensembles tels que $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{D}$ ou même $\mathbb{A}$, sont dénombrables, c'est-à-dire qu'il existe une bijection entre chacun

◆ Base des logarithmes népériens.

d'eux et $\mathbb{N}^*$. Au contraire, l'ensemble des nombres transcendants (et à fortiori celui des nombres réels) est infiniment plus riche que $\mathbb{N}^*$. Comme on le conçoit facilement, il est également dense.
En dépit de ces faits, il est très difficile de démontrer qu'un nombre donné, comme π, est transcendant. On ne connaît explicitement que peu de nombres transcendants ; en déterminer de nouveaux est une performance en théorie des nombres.

Le nombre complexe

Les deux échelons suivants de l'escalade concernent l'ensemble $\mathbb{C}$ des nombres complexes et l'ensemble $\mathbb{H}$ des quaternions ◆, qui sont deux corps. L'ensemble $\mathbb{H}$ des quaternions est une extension de $\mathbb{C}$ aux propriétés particulières. Ainsi ab est-il différent de ba ; l'équation $x^2 = -1$ possède une infinité de racines, etc.
Nous nous limiterons à $\mathbb{C}$, qui correspond à peu près à la généralisation optimale de $\mathbb{N}^*$; il ne faudrait en effet pas croire que toute généralisation est nécessairement uniquement enrichissante ; en passant d'un ensemble à l'autre, on gagne, certes, des propriétés, mais il en est également de perdues. Ainsi, dans $\mathbb{N}^*$, pouvait-on affirmer que l'égalité $ab = 1$ n'était possible que si $a = b = 1$, ce qui est évidemment faux dans $\mathbb{Z}$, $\mathbb{Q}$ ou $\mathbb{R}$. Il faut donc s'arrêter à un certain stade, en décidant de ne plus décerner le label de « nombre » à des éléments de généralisations trop sophistiquées.

◆ L'ensemble des quaternions est l'ensemble des nombres de la forme $a + bi + cj + dk$, où :
1) a, b, c, d sont des nombres réels ;
2) i est le nombre complexe de carré -1 qui intervient dans toute la théorie des nombres complexes ;
3) j et k sont deux nouveaux nombres tels que :
$i \times j = -j \times i = k$;
$j \times k = -k \times j = i$;
$k \times i = -i \times k = j$.
L'ensemble des quaternions, muni de l'addition et de la multiplication, est un corps non commutatif.

On appelle nombre complexe toute écriture de la forme $a + bi$, où a et b sont des nombres réels, i représentant un nouvel élément ; les règles de calcul dans $\mathbb{C}$ sont déterminées par les règles valables dans $\mathbb{R}$, en y ajoutant l'égalité, impossible dans les nombres réels, $i^2 = -1$. Ainsi :

$$(a + ib) + (a' + ib') = (a + a') + i(b + b')$$
$$(a + ib) \times (a' + ib') = (aa' - bb') + i(ab' + a'b)$$

Toute équation a des solutions dans l'ensemble des nombres complexes

A l'origine, le corps $\mathbb{C}$ a été conçu pour pouvoir résoudre toutes les équations du second degré, comme $x^2 + 1 = 0$ ou $x^2 + x + 7 = 0$, qui n'ont aucune racine réelle. On a donc introduit dans $\mathbb{R}$ une sorte de greffon, avec le nombre i tel que $i^2 = -1$. Celui-ci est racine de l'équation $x^2 + 1 = 0$, mais non de $x^2 + x + 7 = 0$; ce qui est remarquable, c'est que celle-ci acquiert quand même des racines qui s'expriment simplement en fonction de i (il s'agit des nombres $\frac{-1 + 3i\sqrt{3}}{2}$ et $\frac{-1 - 3i\sqrt{3}}{2}$) ; ce qui est encore plus remarquable,

c'est que l'introduction du seul nombre i permet de donner une racine à n'importe quelle équation à coefficients réels, et même — ce qui est encore plus surprenant — à coefficients complexes ◆.
En revanche, $\mathbb{C}$ possède une grave lacune par rapport aux autres ensembles de la famille des ensembles numériques ; comme $\mathbb{H}$ d'ailleurs, qui la suit dans la liste :

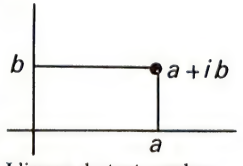
◆ On dit que $\mathbb{C}$ est algébriquement clos ; ce n'est pas le seul corps de ce type ; ainsi existe-t-il une extension de $\mathbb{A}$, faite de nombres complexes particuliers, qui possède la même propriété.

$$\mathbb{N}^* \subset \mathbb{Q}_+^* \subset \mathbb{Q} \subset \mathbb{A} \subset \mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$$
$$\mathbb{Z} \subset \mathbb{D}$$

Dans $\mathbb{C}$, on ne peut définir d'inégalités ayant les mêmes propriétés que dans $\mathbb{R}$; l'existence de i empêche de conserver une « règle des signes » qui assure, notamment, que le produit de deux nombres de même signe (par exemple, égaux) est nécessairement positif, puisque :

$$i^2 = i \times i = 1 < 0.$$

Une représentation commode de $\mathbb{C}$ est donnée par les points du plan : le point de coordonnées a et b est l'image du nombre complexe $a + ib$. Cette représentation fait du corps $\mathbb{C}$ des nombres complexes l'outil idéal pour étudier algébriquement la géométrie du plan.



L'image de tout nombre réel se trouve sur la droite $x'\ x$, l'image des nombres imaginaires purs se trouve sur $y'\ y$. Le nombre complexe est représenté dans le plan.

ETUDE ENSEMBLISTE DES NOMBRES

ENSEMBLES FINIS ET ENSEMBLES INFINIS

Nous supposerons maintenant connues du lecteur définitions et propriétés de la théorie des ensembles ◆. Munis de ces règles, nous allons bâtir une théorie des ensembles finis. Ceux-ci contiennent en effet, nous le sentons bien, toutes les connaissances nécessaires pour définir commodément les entiers. Mais qu'est-ce qu'un ensemble fini ? En parler comme d'un ensemble dont le nombre d'éléments est fini est, au stade où nous sommes, un non-sens, la notion d'ensemble étant évidemment plus générale que celle de nombre. Il nous faut nous résigner à une étude directe de la finitude.
Pour cela, nous partirons d'un proverbe et d'une remarque de Galilée. Le proverbe est : « Le tout est plus grand que la partie. ◆ » Traduit en mathématiques, il signifie quelque chose comme : *Si* F *est un sous-ensemble de* E *différent de* E *lui-même — on dit : sous-ensemble strict —, il n'existe pas de bijection entre* E *et* F ◆.
Galilée fut, semble-t-il, le premier à remarquer que cela est inexact.

◆ Voir l'article « les Ensembles » et les notions de corps et d'anneau dans l'article « l'Algèbre ».

◆ C'est en fait l'un des axiomes d'Euclide.

◆ Voir l'article « les Ensembles », page 262.

S'il est parfaitement vrai que le proverbe se justifie avec les petits ensembles, il devient faux pour de grands ensembles comme $\mathbb{N}$, par exemple (ensembles des entiers positifs, y compris le 0 ; $\mathbb{N} = \{ 0, 1, 2, 3, \ldots \}$. L'application qui associe à tout entier n son carré n^2 est, en effet, une bijection de $\mathbb{N}$ sur l'un de ses sous-ensembles stricts, celui des carrés parfaits.

De la même manière, l'application qui associe à tout entier n son successeur $n' = n + 1$ est une bijection de $\mathbb{N}$ sur l'un de ses sous-ensembles stricts $\mathbb{N}^*$, celui des entiers strictement positifs. De même pour $(n \longmapsto 2n)$, etc.

De telles bijections n'existent évidemment pas pour des ensembles finis, puisqu'un sous-ensemble strict possède un cardinal strictement inférieur à celui de l'ensemble : ce qui interdit toute possibilité de bijection.

Il y a plus intéressant encore : non seulement le proverbe est inexact pour $\mathbb{N}$, mais il est également inexact pour tout ensemble infini (c'est-à-dire non fini), puisqu'un tel ensemble contient toujours au moins un sous-ensemble en bijection avec $\mathbb{N}$ (on dit : un sous-ensemble dénombrable) ◆ : il est alors facile d'y généraliser un exemple du type Galilée.

◆ Voir l'exemple donné dans l'article « les Ensembles », pages 263 et 264.

Notre étude préliminaire, très intuitive certes, nous a donc permis de dégager, avec une grande vraisemblance, un critère essentiel, dû à Dirichlet :

Un ensemble infini est, par définition, un ensemble qui peut être mis en bijection avec l'un de ses sous-ensembles stricts. Un ensemble fini est, par définition, un ensemble qui n'est pas infini.

Un ensemble fini est un ensemble qui est strictement « plus grand » que ses parties strictes ◆.

◆ Cette définition est incomplète, car elle laisse échapper le cas particulier de l'ensemble vide $\emptyset$, qui est bien fini.

Munis de ce critère, nous pouvons bâtir complètement la théorie des ensembles finis. Pour des raisons de commodité, nous partirons, en fait, d'une propriété différente (mise en évidence par Stäckel et Enriques) mais équivalente. Elle procède d'une idée très éloignée, en apparence, mais finalement plus maniable, que nous allons exposer maintenant. Pour ces mathématiciens, le concept fondamental était celui que nous appellerons « comptabilité ». Un ensemble est comptable si l'on peut en énumérer complètement les éléments, par exemple sur une bande magnétique : il existe donc un début et une fin pour cette énumération. De plus, nous remarquons qu'en effaçant çà et là des passages de la bande il nous reste également des ensembles comptables, sous-ensembles du précédent (le plus simple d'entre eux étant l'ensemble vide).

Techniquement, nous prendrons donc comme axiome de départ l'énoncé suivant :

Un ensemble fini est un ensemble E *sur lequel on peut définir une relation d'ordre, notée* ($\leqslant$), *pour laquelle tout sous-ensemble non vide de* E *admet un premier et un dernier élément.*

Il est alors immédiat de vérifier les propriétés suivantes des ensembles finis : si E et F sont finis, si G est un ensemble quelconque, alors tout sous-ensemble de E, l'intersection ($E \cap G$), la réunion ($E \cup F$), l'ensemble $\mathfrak{P}(E)$ des parties de E, tout ensemble qui peut être mis en bijection avec un sous-ensemble de E, le produit cartésien $E \times F$, l'ensemble F^E des applications de E dans F sont des ensembles finis. Pour le démontrer, il suffit de chercher, dans chaque cas, une règle pour énumérer les éléments de l'ensemble étudié et de vérifier que la définition est bien satisfaite.

Nous ne donnerons qu'un seul exemple ; montrons que ($E \cup F$) est fini. Pour cela, nous enregistrerons d'abord les éléments de E avec leur ordre, puis nous compléterons la bande magnétique avec les éléments de F avec leur ordre. Il ne restera plus ensuite qu'à effacer les éléments de F déjà considérés comme éléments de E. Par exemple :

$$E = \{ 1, 4, 8, 9, 15, 27 \}$$

$$F = \{ 4, 6, 7, 11, 15, 16 \}$$

$$E \cup F = \{ 1, 4, 8, 9, 15, 27 \} \cup \{ 4, 6, 7, 11, 15, 16 \}$$

$$= \{ 1, 4, 8, 9, 15, 27, 6, 7, 11, 16 \}$$

C'est un peu plus délicat pour $E \times F$ (ou $\mathfrak{P}(E)$), mais une recherche sur de petits ensembles met rapidement sur la voie.

Avant d'en déduire les propriétés de $\mathbb{N}$, nous ne pouvons pas manquer de signaler que la définition que nous exploitons ici n'est autre qu'un raffinement d'une notion tout à fait fondamentale en mathématiques : celle d'ensemble bien ordonné, c'est-à-dire d'ensemble muni d'une relation d'ordre pour laquelle tout sous-ensemble non vide admet un premier élément ◆. Un ensemble fini est donc un ensemble doublement bien ordonné (P. Dubreil), c'est-à-dire un ensemble bien ordonné pour la relation $\leqslant$ et pour la relation opposée $\geqslant$.

◆ Voir l'article « les Ensembles », page 282.

Nous sommes maintenant en possession d'un outil très commode ; reste à extraire, grâce à lui, la notion de nombre entier de celle d'ensemble fini. Ici est nécessaire un acte de foi, c'est-à-dire un nouvel axiome, qui fait partie des axiomes de la théorie des ensembles usuelle ◆. Des ensembles qui peuvent être reliés par des bijections forment une classe ◆ — approximativement, un « gros » ensemble, si gros que des paradoxes apparaîtraient si on le considérait comme un ensemble ; mais c'est néanmoins un objet mathématique. Dans chacune des classes ainsi formées, choisissons un ensemble particulier, appelé cardinal de la classe ou encore cardinal de chacun des ensembles de la classe.

◆ Il apparaît notamment, sous plusieurs formes équivalentes, dans l'axiomatique de Zermelo-Fraenkel.

◆ Voir l'exemple donné dans l'article « les Ensembles », page 262 : « L'ensemble de tous les ensembles n'existe pas. »

L'ENSEMBLE DES NOMBRES ENTIERS NATURELS : $\mathbb{N}$

L'acte de foi dont nous parlions consiste en ceci : nous admettrons que les cardinaux des ensembles finis (on dit : cardinaux finis) forment un ensemble. Cet ensemble n'est autre que $\mathbb{N}$; les cardinaux finis sont donc les entiers naturels. 0 est le cardinal de la classe de l'ensemble vide $\varnothing$, 1 celui de la classe de l'ensemble $\{\varnothing\}$, 2 celui de la classe de l'ensemble $\{\varnothing, \{\varnothing\}\}$.

La construction explicite de von Neumann ◆ — qui n'est qu'une construction parmi d'autres — consiste donc à poser, par définition :

$$0 = \varnothing,\ 1 = \{\varnothing\} = \{0\} = \{a\},\ 2 = \{\varnothing, \{\varnothing\}\} = \{0, 1\},$$
$$n' = \{0, 1, 2, \ldots, n\}\ (= n + 1).$$

◆ J. von Neumann (1903-1957) : mathématicien américain qui a créé une des équations fondamentales de la physique théorique, fondé les relations entre cybernétique et ordinateurs, et fourni les principes de base de la programmation.

Nous lisons : 0 est le cardinal de l'ensemble ne contenant aucun élément ; 1 est le cardinal de l'ensemble contenant 1 élément (ici, l'ensemble vide) ou l'élément 0, ou l'élément a ; 2 est le cardinal de l'ensemble contenant 2 éléments, l'ensemble vide et l'ensemble contenant l'ensemble vide, ou 0 et 1.

$\mathbb{N}$ étant défini, il reste naturellement à en déduire les propriétés fondamentales. Nous n'indiquerons que très sommairement celles-ci, nous contentant de les énumérer dans l'ordre le plus simple possible. Tout d'abord, $\mathbb{N}$ est ordonné par la relation suivante : « $n \leqslant m$ si et seulement s'il existe une injection de n dans m » (n'oublions pas que n et m sont des ensembles !). Le segment $[0, m]$ étant l'ensemble des entiers n tels que $(n \leqslant m)$, on peut montrer qu'il est fini ; il suffit pour cela d'utiliser la relation d'injectibilité définie ci-dessus. Tout sous-ensemble non vide de $\mathbb{N}$ admet un premier élément (0 est le premier élément de $\mathbb{N}$ lui-même) ; tout sous-ensemble non vide de $\mathbb{N}$ qui est majoré (c'est-à-dire inclus dans un segment $[0, m]$) admet un dernier élément.

Nous sommes presque arrivés au bout de notre construction, mais il manque encore un résultat essentiel. $\mathbb{N}$ n'admet pas de dernier élément : *il n'existe pas d'entier supérieur à tous les autres* ◆.

◆ Ce résultat peut être démontré par le théorème de Cantor (il n'existe aucune injection entre un ensemble $\mathcal{P}(E)$ et l'ensemble E). Voir l'article « les Ensembles », page 284.

De cela découle la dernière propriété fondamentale : $\mathbb{N}$ est infini. On peut en dégager alors le principe de récurrence, sans lequel de très nombreuses propriétés mathématiques resteraient inaccessibles. C'est la formalisation d'une remarque très naturelle, suivant laquelle tout entier peut être obtenu par un « nombre fini » d'étapes à partir du minorant absolu 0. Intuitivement, nous savons tous ce qu'est le successeur du nombre n : c'est l'entier n', assimilé par la suite à $(n + 1)$, qui vient « juste après » n. Dans notre système, n' est simplement le premier des éléments qui ne sont pas inférieurs à n. Le principe de récurrence utilise pleinement cette notion très simple ; techniquement, on l'écrit sous la forme suivante :

Tout sous-ensemble de $\mathbb{N}$ *qui contient* 0 *et les successeurs de tous ses éléments est confondu avec* $\mathbb{N}$.

Cela met un point final à notre recherche d'une caractérisation rationnelle, uniquement fondée sur la théorie des ensembles, des nombres entiers naturels à partir des ensembles finis.
Nous avons insisté longuement sur la possibilité d'une théorie presque totalement élémentaire de $\mathbb{N}$; seuls deux ou trois résultats, comme celui qui établit l'existence d'une injection reliant deux ensembles donnés à priori, sont difficiles et supposent l'axiome du choix. L'exposé minutieux de son déroulement permet à chacun de se persuader que la définition et la description de $\mathbb{N}$ ont pu être totalement dégagées de tout recours, implicite ou non, à l'intuition qui nous vient de notre manipulation quotidienne des entiers. On y voit, par exemple, que l'existence d'un plus petit entier possédant une propriété donnée est effectivement démontrée ; une telle existence (utilisée si souvent par Fermat, par exemple ◆), bien qu'évidente, repose en fait sur des relations très profondes entre ensembles.
Enfin, cet exposé nous a permis d'exhiber plusieurs propositions sur $\mathbb{N}$ dont l'importance pratique est primordiale. Nous en ferons notamment deux groupes :

◆ En théorie des nombres, pour démontrer qu'il n'existe aucun entier satisfaisant à une certaine propriété, on fait souvent un raisonnement par l'absurde où intervient le plus petit entier ayant cette propriété.

Le premier est formé des trois théorèmes :
a) $\mathbb{N}$ est muni d'une relation d'ordre ($\leqslant$) pour laquelle tout sous-ensemble non vide admet un premier élément ($\mathbb{N}$ est bien ordonné) ;
b) tout sous-ensemble non vide majoré admet un dernier élément ;
c) $\mathbb{N}$ n'est pas majoré.

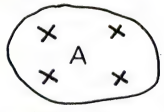
Le second est formé du principe de récurrence, joint aux deux théorèmes :
a) il existe dans $\mathbb{N}$ une injection qui associe à tout entier n son successeur n' ;
b) il existe dans $\mathbb{N}$ un élément noté 0 qui n'est le successeur d'aucun entier.

Maintenant que nous savons que les fondations de $\mathbb{N}$ reposent sur la grande théorie des ensembles, nous pouvons couper le cordon ombilical et ne garder que l'un ou l'autre de ces systèmes d'axiomes : c'est souvent le second qui est choisi ; il porte le nom de système de Dedekind-Peano (1889) ◆.

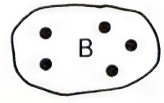
◆ R. Dedekind (1831-1916) : mathématicien allemand qui fut le premier à analyser avec rigueur le concept de nombre irrationnel.
G. Peano (1858-1932) : mathématicien italien qui inventa un langage mathématique universel, tout en symboles.

De ces axiomes découle la théorie des ensembles finis (c'est la démarche inverse de celle que nous avons suivie ici) ; on peut définir l'addition et la multiplication à partir de la réunion et du produit ensemblistes.

Prenons deux ensembles finis A et B de cardinaux respectifs m et n et tels que l'intersection de A et B soit vide :

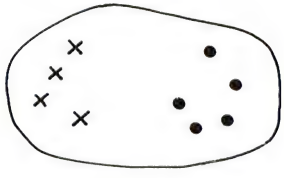


Card A = 4 = m



Card B = 5 = n

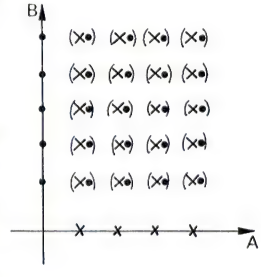
Nous appelons « somme » de m et n : $m + n$ le cardinal de l'ensemble A U B.



p = Card A U B = $m + n$

Nous appelons « produit » de m et n : $m \times n$ le cardinal de l'ensemble-produit A × B (ensemble des couples) ◆ (voir schéma ci-contre).

◆ Voir l'article « les Ensembles », pages 268, 269.



On peut également définir ces opérations directement, par les égalités :

$$\begin{cases} n + 0 = n \\ n + m' = (n + m)' \end{cases} \qquad \begin{cases} n \times 0 = 0 \\ n \times m' = (n \times m) + n \end{cases}$$

Si certaines démonstrations sont un peu longues et reposent sur un grand nombre de récurrences successives, elles ne sont pas difficiles.

L'étude de la relation d'ordre ($n \leqslant m$ si et seulement si on peut trouver p tel que $n + p = m$) est également facile à conduire. Très vite, on est amené aux portes de l'arithmétique traditionnelle, dès lors solidement fondée, dans laquelle l'égalité classique (2 + 2 = 4) possède enfin une justification !

L'étude des entiers débouche sur des problèmes complexes où interviennent principalement les notions de divisibilité, de nombre premier, de p.g.c.d. et de p.p.c.m. ◆.

◆ Le p.g.c.d. (plus grand commun diviseur) et le p.p.c.m. (plus petit commun multiple) sont des notions fondamentales en arithmétique. Leur nom même suffit à les définir.

Certains nombres entiers sont divisibles par d'autres entiers. Par exemple, 24 = 3 × 8 = 2 × 12 = 4 × 6... Par contre, d'autres, comme 1, 2, 3, 7, 11, 13, 127..., n'admettent pas d'autres diviseurs que 1 et eux-mêmes. Ce sont des *nombres premiers*.

L'ENSEMBLE DES ENTIERS RELATIFS : ℤ

Revenons maintenant à l'extension de ℕ qui nous a conduits à la notion d'entier relatif, c'est-à-dire — grosso modo — d'entier muni d'un signe.

Rappelons brièvement que la soustraction n'est pas toujours possible, mais que, si cela a un sens, les soustractions

$$a - b = x,\ c - d = x$$

donnent le même résultat si et seulement si on peut écrire :

$$a + d = b + c$$

Ainsi a-t-on défini des couples équivalents en les liant par cette relation : soustraire b de a équivaut à soustraire d de c si et seulement si $a + d = b + c$.

Franchissons un pas en supprimant l'hypothèse de l'existence de l'entier x, différence de a et de b. Nous pouvons encore écrire l'équivalence logique :

$$(a - b = c - d) \iff (a + d = b + c).$$

Par exemple, $2 - 7 = 4 - 9$, puisque $2 + 9 = 7 + 4$. La différence est qu'ici le symbole $a - b$, employé seul, n'a pas de sens ; ce n'est que dans le cadre d'une égalité $(a - b = c - d)$ qu'il peut être utilisé. Dès lors, la démarche est évidente. Le nombre x, qui, ajouté à 5, donnerait 0, n'existe pas. S'il existait, on pourrait écrire :

$$x = 0 - 5 = 1 - 6 = 2 - 7 = 3 - 8 = 4 - 9 = \ldots,$$

mais ces égalités n'ont pas de sens. Décidons donc de leur donner, d'un coup, un sens : le nouveau nombre ainsi créé sera noté (-5), abréviation de $(0 - 5)$. Techniquement, c'est l'ensemble des couples ◆ (a, b) liés par $a + 5 = b$.

◆ Voir l'article « les Ensembles », page 268.

Le nombre $(+5)$ peut être défini de manière tout analogue ; il suffit d'écrire :

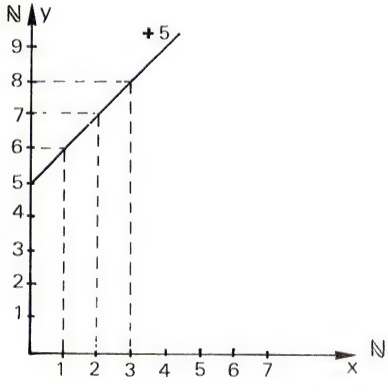
$$+5 = 5 - 0 = 6 - 1 = 7 - 2 = 8 - 3 = 9 - 4 = \ldots$$

Il y aurait peu d'intérêt, dans la pratique, à distinguer trop soigneusement le naturel 5, élément de $\mathbb{N}$, et le positif

$$(+5) = \{ (5, 0), (6, 1), (7, 2), \ldots \},$$

sous-ensemble du produit cartésien ◆ $\mathbb{N} \times \mathbb{N}$: on les confond donc toujours.

◆ Voir l'article « les Ensembles », page 269.

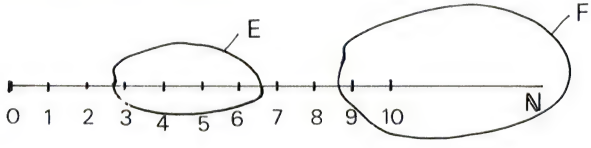


Les points de cette droite représentent le nombre $+5$. C'est un sous-ensemble du produit cartésien $\mathbb{N} \times \mathbb{N}$ formés des couples (y, x) vérifiant la propriété $y - x = 5$.

L'ensemble des nombres que nous venons de définir ainsi se note $\mathbb{Z}$ (de l'allemand « Zahlen », nombre) :

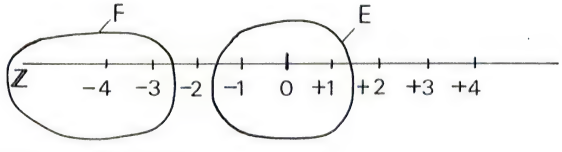
$$\mathbb{Z} = \{ \ldots, -3, -2, -1, 0, 1, 2, 3, 4, 5, \ldots \},$$

et s'appelle l'ensemble des entiers relatifs. L'ensemble des entiers positifs ou naturels $\mathbb{N}$ est un sous-ensemble de l'ensemble des entiers relatifs (positifs et négatifs) $\mathbb{Z}$. On peut se poser la question : les propriétés de $\mathbb{N}$ peuvent-elles être étendues à $\mathbb{Z}$? Il est possible de définir une addition dans $\mathbb{Z}$. La relation d'ordre définie dans $\mathbb{N}$ peut s'étendre à $\mathbb{Z}$, mais avec d'autres propriétés. Par exemple, il n'est pas toujours possible de trouver un plus petit élément d'un sous-ensemble donné de $\mathbb{Z}$, alors que, dans $\mathbb{N}$, il est toujours possible de trouver un plus petit élément :



E a comme plus petit élément 3 ◆

F a comme plus petit élément 9



E a comme plus petit élément — 1

F n'a pas de plus petit élément

Extension d'un ensemble par passage au quotient

Partons d'un ensemble E muni d'une certaine structure, par exemple $E = \mathbb{N}$, muni de l'addition. Admettons qu'il soit possible de regrouper tous les éléments de $E \times E$ (ensemble des couples (a, b) où a et b appartiennent à E) en classes, deux à deux sans élément commun, telles que la somme d'un couple quelconque pris dans une classe α et d'un couple quelconque pris dans une classe β donne toujours un couple d'une même classe γ. Nous dirons alors que l'on a défini, sur l'ensemble E′ (appelé quotient), une addition analogue à celle de E, en posant $\alpha + \beta = \gamma$. Souvent, le nouvel ensemble E′ possède un sous-ensemble presque identique — on dit : isomorphe — à E. On dit alors que E′ est une extension de E. Dans notre exemple, $E = \mathbb{N}$. (a, b) et (c, d) seront dans la même classe si et seulement si $a + d = b + c$.

Alors l'ensemble quotient E′ n'est autre que $\mathbb{Z}$. Si l'on confond, par commodité, l'entier n et la classe α du couple $(n, 0)$, on peut alors écrire $\mathbb{N} \subset \mathbb{Z}$.

Comme nous verrons plus loin que les corps classiques en mathématiques s'obtiennent également à partir des entiers (corps des rationnels, des nombres algébriques, des réels, des complexes, des quaternions, des nombres p-adiques), $\mathbb{Z}$ est donc à l'origine de toute la théorie des corps usuels.

L'ENSEMBLE DES NOMBRES RATIONNELS

Il nous faut maintenant quitter les ensembles $\mathbb{N}$ et $\mathbb{Z}$ pour entrer dans des familles de nombres plus régulières, où les problèmes posés sont d'une nature différente. Nous étions dans le domaine du discret, du fini, des dénombrements ; les extensions successives sont, à partir de maintenant, beaucoup plus liées à la notion de continu, d'infini, de mesure (intégration).
De même que l'anneau $\mathbb{Z}$ avait été construit pour que la soustraction ($x = a - b$) soit toujours possible, $\mathbb{Q}$ est l'extension de $\mathbb{Z}$ dans laquelle la division ($x = a : b$) est toujours possible, à la seule exception de la division par 0. Le schéma est approximativement le même ; comme la relation :

$$x = \frac{a}{b} = \frac{c}{d}$$

implique l'égalité ($ad = bc$), nous regrouperons ensemble les couples (a, b), avec $b \neq 0$, d'après cette relation d'équivalence. Par exemple, nous écrirons :

$$\frac{12}{28} = \frac{-6}{-14} = \frac{36}{84} = \frac{-3}{-7} = \frac{-24}{-56} = \ldots,$$

bien que ces symboles, pris isolément, n'aient pas de sens. Leur en donner un constitue précisément l'extension de $\mathbb{Z}$ à $\mathbb{Q}$; le nombre rationnel $\frac{3}{7}$ est, en fait, la classe de tous les couples (ou fractions) équivalant à $\frac{3}{7}$, comme $\frac{12}{28}, \frac{-6}{-14}$, etc. Là encore, il faut définir dans $\mathbb{Q}$ des opérations qui se raccordent avec celles de $\mathbb{Z}$.
Naturellement, si cette extension conserve la plupart des propriétés de $\mathbb{Z}$, elle ne saurait les garder toutes. Il en était de même pour le passage de $\mathbb{N}$ à $\mathbb{Z}$; dans le premier ensemble $\mathbb{N}$, l'égalité ($x + y = 0$) est équivalente aux deux conditions ($x = 0$, $y = 0$), ce qui n'est plus exact dans $\mathbb{Z}$, où ($3 + (-3) = 0$). De même, l'égalité ($xy = 1$) dans $\mathbb{Z}$ entraîne les égalités ($x^2 = 1$, $y^2 = 1$), ce qui est faux dans $\mathbb{Q}$, où $\left(\frac{2}{3} \times \frac{3}{2} = 1\right)$. Une extension n'est utile que si le rapport des propriétés conservées à celles qui disparaissent est appréciable : c'est le problème de toutes les généralisations.

$\mathbb{Q}$ ensemble des rationnels est donc formé des entiers positifs, négatifs et de toutes les fractions possibles d'entiers. $\mathbb{Q}$ est un corps ; mais il existe, entre $\mathbb{Z}$ et $\mathbb{Q}$, des ensembles intéressants, comme l'anneau $\mathbb{D}$ des nombres décimaux ◆. Ce sont des nombres fractionnaires qui peuvent être représentés par des fractions ayant en dénominateur des puissances de 10 et, par conséquent, le dénominateur réduit ne contient que des 2 et des 5. Par exemple :

◆ Voir ci-dessus page 371 et lire Décimal (nombres).

$$\frac{28}{100} = \frac{7 \times 2 \times 2}{5 \times 5 \times 2 \times 2} = \frac{7 \times 2}{5 \times 5 \times 2} = \frac{7}{5 \times 5}$$

Les tout premiers temps de l'introduction historique des nombres rationnels eux-mêmes virent l'emploi à peu près exclusif d'inverses $\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right)$ et de fractions très simples $\left(\frac{2}{3}\right)$; il fallut disposer d'une numération commode — notre numération *décimale* — pour pouvoir effectuer réellement des calculs précis sur les rationnels ; c'est la notation « à virgule » qui permet cela. On y remplace $\frac{2}{5}$ par 0,4, etc.

La difficulté majeure provient de ce que l'anneau $\mathbb{D}$ des nombres décimaux ne recouvre pas $\mathbb{Q}$ tout entier ; le nombre $\frac{1}{3}$ ne peut s'y exprimer que par une « expansion » infinie de décimales : 0,333 333... Tout rationnel non décimal admet ainsi un développement infini qui est périodique au-delà d'un certain rang. Ainsi :

$$\frac{19}{28} = 0{,}67\ 857142\ 857142\ 857142\ldots$$

(signalons au passage que la considération de ces périodes conduit à de passionnants et difficiles problèmes... de théorie des nombres). Par le biais de l'approximation, indispensable, de rationnels par des décimaux, biais justifié par la commodité des calculs, on est donc habitué à concevoir qu'un nombre peut très bien être défini par la donnée de nombres décimaux successifs, encadrant de mieux en mieux le nombre lui-même. Ainsi :

$$0 < \frac{19}{28} < 1\ ; \quad 0{,}6 < \frac{19}{28} < 0{,}7\ ; \quad 0{,}67 < \frac{19}{28} < 0{,}68\ ;$$

$$0{,}678 < \frac{19}{28} < 0{,}679\ ; \quad 0{,}678\ 5 < \frac{19}{28} < 0{,}678\ 6\ ;$$

$$0{,}678\ 57 < \frac{19}{28} < 0{,}678\ 58, \text{ etc.}$$

A première vue, rien ne distingue deux développements décimaux comme :

0,678 571 428 ... et 3,141 592 653 ...

Il existe des rationnels de période décimale excessivement longue, et personne ne peut prouver, à la seule vue des encadrements successifs du nombre π, que ses approximations décimales ne finiront pas par être périodiques.

L'hypothèse suivant laquelle le nombre π, longueur d'un cercle de diamètre 1, existerait — c'est-à-dire serait rationnel, au point où nous en sommes — permet des calculs se traduisant par des encadrements décimaux :

$3 < \pi < 4$; $3{,}1 < \pi < 3{,}2$; $3{,}14 < \pi < 3{,}15$;

$3{,}141 < \pi < 3{,}142$; $3{,}141\ 5 < \pi < 3{,}141\ 6$;

$3{,}141\ 59 < \pi < 3{,}141\ 60$, etc.

L'ENSEMBLE DES NOMBRES IRRATIONNELS

Nous sommes prêts, maintenant, pour le saut décisif : pour nous, désormais, tout développement décimal définit un nombre ; s'il est périodique, ce nombre est rationnel; par exemple :

$x = 0{,}678\ 571\ 428\ 571\ 42 \ldots = 0{,}67 + 0{,}857\ 142\ y$,

avec $y = \dfrac{1}{100} \times (1{,}000\ 001\ 000\ 001\ 000\ 001\ 00\ldots)$,

d'où $999\ 999\ y = \dfrac{1}{100} \times (999\ 999{,}\ 999\ 999\ 999\ 9\ldots) = 10\ 000$, d'où

$$x = \frac{67}{100} + \frac{1}{100} \times \frac{857\ 142}{999\ 999} = \frac{67}{100} + \frac{6}{700} = \frac{475}{700} = \frac{19}{28}.$$

Les développements non périodiques définissent des irrationnels. $\sqrt{2}$ est un tel nombre (π également, bien que ce ne soit pas évident du tout). Tel est, approximativement, le déroulement du processus à demi inconscient qui a permis dans le passé (et qui permet encore aux enfants qui prennent contact pour la première fois avec les nombres réels) d'admettre parfaitement que l'on comble ainsi les « trous » de l'ensemble des nombres rationnels. Nous verrons que la définition correcte des réels diffère de ce schéma intuitif — bien que l'on puisse, mal commodément, le rendre rigoureux —, mais qu'elle justifie, après coup, cette audace somme toute limitée.

L'ENSEMBLE DES NOMBRES RÉELS

Une définition précise du corps des nombres réels est assez délicate. Plusieurs voies sont possibles. Il s'agit de « boucher des trous » existant dans $\mathbb{Q}$. Ainsi, un nombre positif x tel que $x^2 = 2$ n'existe

pas dans $\mathbb{Q}$; s'il existait, on pourrait pourtant en obtenir des encadrements, puisque, par exemple :

$(1{,}41)^2 = 1{,}988\ 1 < 2 < 2{,}016\ 4 = (1{,}42)^2$,

d'où $\quad 1{,}41 < x < 1{,}42$

On pourrait théoriquement découvrir de cette façon des encadrements de plus en plus serrés de ce nombre... inexistant.

L'une des façons de procéder, qui n'est simple qu'en apparence, consiste à définir d'abord les réels compris entre 0 et 1. On y retrouve tous les décimaux et même tous les rationnels compris entre les mêmes limites ; par exemple, on y reconnaît :

$$\frac{19}{28} = 0{,}678\ 571\ 428\ 571\ 428\ 571\ 42\ \ldots$$

Il est alors facile d'avoir tous les nombres réels, en ajoutant à un nombre entier relatif arbitraire — appelé partie entière — un nombre réel x tel que $0 \leqslant x < 1$ (x est la mantisse du réel).

La définition de la somme et du produit de deux nombres réels est un peu délicate, elle aussi. Disons grossièrement que l'on remplace, pour les calculer, les nombres réels par des nombres décimaux obtenus en ne gardant que les premiers chiffres de chaque développement ; plus on garde de chiffres, plus on obtient une somme et un produit approchant la véritable somme et le véritable produit.

Ainsi, sachant que :

$\sqrt{2} = 1{,}414\ 2\ldots$, $\sqrt{3} = 1{,}732\ 0\ldots$,

on voit que $\sqrt{2} \times \sqrt{3}$ n'est pas très différent de :

$$1{,}41 \times 1{,}73 = 2{,}439\ 3$$

(en fait, $\sqrt{2} \times \sqrt{3} = \sqrt{6} = 2{,}449\ 5\ldots$).

Tout cela permet de munir $\mathbb{R}$, l'ensemble des nombres réels, d'une structure de corps, voisine de celle de $\mathbb{Q}$, l'ensemble des nombres rationnels. $\mathbb{R}$ possède de très nombreuses propriétés ; c'est l'instrument de travail nécessaire à tous, notamment aux physiciens.

Signalons quelques-unes de ces propriétés :

— Entre deux réels distincts, il existe une infinité de nombres réels, dont une infinité de rationnels (éléments de $\mathbb{Q}$) et une infinité d'irrationnels.

— Si un polynôme réel p est tel que

$$a < b, \qquad p(a) < 0 < p(b)$$

il existe au moins un réel x tel que

$$a < x < b \qquad \text{et} \qquad p(x) = 0$$

(la propriété analogue est fausse dans $\mathbb{Q}$, comme on le voit en prenant $p(x) = x^2 - 2$, $a = 0$ et $b = 2$).

— Pour tout ensemble E non vide de nombres réels tous inférieurs à un nombre a, il existe un réel b tel que :
a) si x appartient à E, alors $x \leqslant b$;
b) si $y < b$, il existe un élément x de E tel que $y < x \leqslant b$.
— Il n'existe pas de bijection entre $\mathbb{R}$ et $\mathbb{N}$ ni entre $\mathbb{R}$ et $\mathbb{Q}$, bien que, en dépit des apparences, il existe des bijections entre $\mathbb{Q}$, $\mathbb{Z}$ et $\mathbb{N}$. Par contre, il existe une bijection entre $\mathbb{R}$ et $\mathcal{P}(\mathbb{N})$, l'ensemble des parties de $\mathbb{N}$.

L'analyse, avec ses délicates questions de limite (d'où les dérivées, les intégrales, etc.), n'est possible que dans $\mathbb{R}$, car il existe trop de « trous » dans l'ensemble des nombres rationnels ◆.

◆ Voir dans l'article « Analyse » la partie consacrée à l'étude de l'ensemble des nombres $\mathbb{R}$, page 101.

L'ENSEMBLE DES NOMBRES COMPLEXES

Les nombres réels sont nécessaires pour pouvoir donner une abscisse à tous les points d'une droite, c'est-à-dire pour pouvoir repérer tous les points de celle-ci.
Il faut deux nombres, réels tous les deux, une abscisse et une ordonnée, pour repérer tous les points d'un plan. Pour certaines questions, il est commode de grouper ces deux coordonnées en un couple ; au point d'abscisse x et d'ordonnée y, on associe ainsi le couple (x, y).
Une autre notation commode consiste à écrire :

$$(x, y) = x + iy.$$

La lettre i est, quant à elle, un pur symbole, comme le signe $+$. Pourtant, les mathématiciens n'ont pas choisi cette notation au hasard. Il serait assez facile, mais trop long, d'expliquer comment ils en sont venus à munir l'ensemble des couples (x, y), appelés nombres complexes, d'une addition et d'une multiplication :

écriture en couples :	écriture avec i :
$(x, y) + (x', y') = (x + x', y + y')$	$(x + iy) + (x' + iy') = (x + x') + i(y + y')$
$(x, y) \times (x', y') = (xx' - yy', xy' + x'y)$	$(x + iy) \times (x' + iy') = (xx' - yy') + i(xy' + x'y)$

Cette introduction a de nombreuses conséquences.
En particulier :

$$x + iy = (x, 0) + [(0, 1) \times (y, 0)]$$

Il est d'usage de confondre volontairement x et $(x, 0)$, y et $(y, 0)$, i et $(0, 1)$.
Les réels x et y sont donc des cas particuliers de nombres complexes.

De même, i est un nombre complexe. Son carré est égal à (-1), puisque :

$$i^2 = i \times i = (0 + 1i) \times (0 + 1i) = -1 + i0 = -1$$

Aucun nombre réel ne possède cette propriété.
L'ensemble $\mathbb{C}$ des nombres complexes, qui contient donc $\mathbb{R}$ comme sous-ensemble, est lui aussi un corps.

Notons la suite d'inclusions :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Les trois derniers ensembles sont des corps, les deux précédents des anneaux.
Ce corps partage avec $\mathbb{Q}$ et $\mathbb{R}$ de nombreuses propriétés, mais il ne peut être question, contrairement à $\mathbb{Z}$, $\mathbb{D}$, $\mathbb{Q}$ ou $\mathbb{R}$, de le munir d'une relation d'ordre, c'est-à-dire de le partager en deux sous-ensembles de nombres, appelés positifs et négatifs, tels que la somme et le produit de deux nombres positifs soient encore positifs. Cette différence est capitale.
Par contre, il possède une propriété absolument essentielle (bien que nous ne puissions expliquer toute la portée de ce résultat) : pour tout polynôme à coefficients complexes — ou réels, en particulier —, on peut écrire une décomposition très simple où la résolution de l'équation $p(x) = 0$ se ramène à des équations du premier degré. Ainsi :

$$x^3 - 2x^2 + x - 2 = (x - 2)(x + i)(x - i)$$

Ce qui montre que l'équation

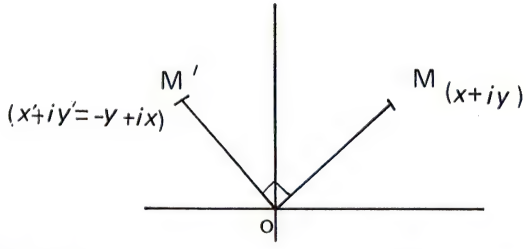
$$x^3 - 2x^2 + x - 2 = 0$$

possède exactement trois racines dans $\mathbb{C}$:

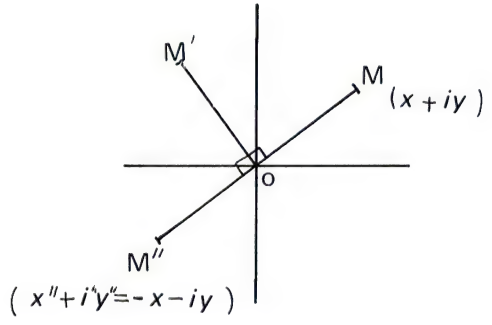
$$x = 2, \qquad x = i, \qquad x = -i$$

Notons qu'elle n'en possède qu'une dans $\mathbb{R}(x = 2)$.
Ce théorème, connu sous le nom de d'Alembert, est dû à Gauss ; on l'appelait autrefois « théorème fondamental de l'algèbre ».
L'analyse définie dans $\mathbb{C}$ est très riche, aussi riche que celle de $\mathbb{R}$; il y existe des dérivées, des intégrales, etc.

Les nombres complexes ont de nombreuses applications, par exemple en géométrie plane (d'où une utilisation constante en physique, pour étudier le courant alternatif par exemple). Pour donner une idée de ces applications, disons simplement que multiplier par i le nombre $x + iy$ revient à remplacer le point M de coordonnées (x, y) par le point M′ image de M dans la rotation d'angle droit :



On comprend dès lors la relation $i^2 = -1$: elle signifie que l'exécution de deux rotations successives d'angle droit revient à une multiplication par (— 1), c'est-à-dire à remplacer M par le point M'' tel que 0 soit le milieu de MM'' (le passage de M à M'' est la symétrie de centre 0) :



Bien que le calcul matriciel ait diminué cette utilisation des nombres complexes en géométrie (domaine un peu délaissé aujourd'hui), les nombres complexes restent essentiels en mathématiques ; sans $\mathbb{R}$ et $\mathbb{C}$, on ne peut comprendre grand-chose à une foule de propriétés parfois très éloignées en apparence de ces notions. Sans pouvoir insister ici, car les difficultés techniques seraient grandes, signalons par exemple un résultat important de théorie des nombres, dû à Dirichlet, selon lequel il existe une infinité de nombres premiers p de la forme :

$$p = an + b$$

si et seulement si 1 est le seul diviseur commun aux nombres entiers a et b. Ce théorème, où n'interviennent à priori que des éléments de $\mathbb{N}$, ne peut encore aujourd'hui être démontré que grâce à des théories délicates où intervient le corps $\mathbb{C}$ des nombres complexes.

André Warusfel.

NORMALE pages 494 à 505.

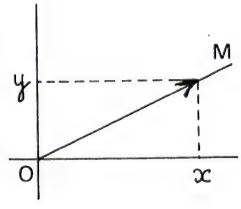
A une courbe. Si la courbe plane (C) admet une tangente au point A, on appelle normale à (C) la droite perpendiculaire en A à la tangente.

A une surface. Si la surface (S) admet au point A un plan tangent, la normale à (S) en A est la droite perpendiculaire en A au plan tangent.

NORME, NORMÉ◆ pages 95, 438.

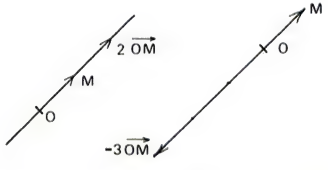
◆ Lire le début de l'article « l'Algèbre linéaire ».

A chaque vecteur du plan, d'origine O et d'extrémité M, associons sa longueur (qui est un nombre réel positif ou nul) que l'on va noter $\| \overrightarrow{OM} \|$ (et que l'on peut calculer, selon le théorème de Pythagore, par la formule $\| \overrightarrow{OM} \| = \sqrt{x^2 + y^2}$ si x et y sont les composantes de $\overrightarrow{OM}$) et examinons-en les propriétés.



— $\| \overrightarrow{OM} \|$ n'est nulle que si O et M sont confondus, c'est-à-dire si le vecteur $\overrightarrow{OM}$ est le vecteur nul $\vec{0}$:

$$\| \overrightarrow{OM} \| = 0 \Longleftrightarrow \overrightarrow{OM} = \vec{0}.$$

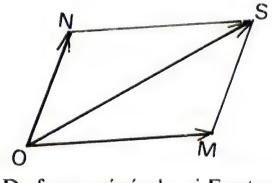


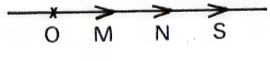
— Si on multiplie le vecteur $\overrightarrow{OM}$ par un nombre réel quelconque λ, la longueur de $\overrightarrow{OM}$ est multipliée par $|\lambda|$ (si on multiplie $\overrightarrow{OM}$ par 2 ou — 3, sa longueur est multipliée par 2 ou 3).

donc : $\forall \| \lambda \overrightarrow{OM} \| = |\lambda| . \| \overrightarrow{OM} \|$.

— Si on considère deux vecteurs $\overrightarrow{OM}$ et $\overrightarrow{ON}$, la longueur du vecteur $\overrightarrow{OS}$ qui est la somme des deux vecteurs $\overrightarrow{OM}$ et $\overrightarrow{ON}$ est inférieure ou égale à la somme des longueurs des deux vecteurs $\overrightarrow{OM}$ et $\overrightarrow{ON}$ (elle n'est égale à cette somme que si O, M, N sont alignés), donc :

$$\| \overrightarrow{OM} + \overrightarrow{ON} \| \leqslant \| \overrightarrow{OM} \| + \| \overrightarrow{ON} \|.$$





De façon générale, si E est un espace vectoriel sur $\mathbb{R}$ (ensemble des nombres réels), on dira qu'on a défini une norme sur E si on a défini sur E une application qui à tout élément x de E fait correspondre un nombre réel positif ou nul noté $\| x \|$ (on lit : norme de x) et qui satisfait aux trois propriétés :

1) La norme de x est nulle si et seulement si x lui-même est nul :

$$\| x \| = 0 \Longleftrightarrow x = 0.$$

2) Quel que soit le nombre réel λ, si on multiplie x par λ, la norme de $\lambda\ x$ est égale au produit de la norme de x par la valeur absolue de λ :

$$|| \lambda . x || = | \lambda | . || x ||.$$

3) Quels que soient les éléments x et y de E, la norme de la somme $x + y$ est inférieure ou égale à la somme des normes de x et de y :

$$|| x + y || \leqslant || x || + || y ||.$$

Un espace vectoriel sur lequel est définie une norme est appelé un espace vectoriel normé.

On a vu plus haut que l'espace vectoriel des vecteurs d'origine O était un espace vectoriel normé, la norme étant l'application qui au vecteur $\overrightarrow{OM}$ de composantes (x, y) associe le nombre réel positif $\sqrt{x^2 + y^2}$.

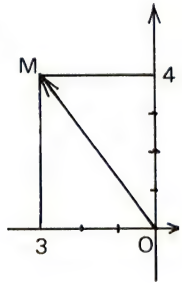
On peut vérifier que les deux applications ci-dessous sont aussi des normes sur l'espace vectoriel des vecteurs d'origine O :

— celle qui au vecteur $\overrightarrow{OM}$ de composantes (x, y) associe le nombre réel positif sup. $(| x |, | y |)$, c'est-à-dire le plus grand des deux nombres positifs $| x |, | y |$;

— ou encore celle qui à $\overrightarrow{OM}$, de composantes (x, y), associe le nombre réel positif $| x | + | y |$.

Calculons, pour un même vecteur $\overrightarrow{OM}$, la norme de ce vecteur selon que l'on a choisi la première, la seconde ou la troisième formule.

Prenons $\overrightarrow{OM}$ de composantes $(-3, +4)$:



— la 1re norme :

$$\sqrt{(-3)^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5 ;$$

— la 2e norme :

$$\sup (| -3 |, | 4 |) = \sup (3, 4) = 4 ;$$

— la 3e norme :

$$| -3 | + | 4 | = 3 + 4 = 7.$$

NOYAU : voir l'article « l'Algèbre », page 43 et l'article « l'Algèbre linéaire », page 80.

NUMÉRATION pages 170, 372.

Le système de numération que nous utilisons est le système décimal, mais il n'est qu'un cas particulier de ce qu'on appelle les systèmes de numération de base a (a étant un entier naturel au moins égal à deux).

Un système de numération de base a comporte a symboles, les chiffres◆ du système (le système à base deux, ou binaire, comporte deux chiffres : 0 et 1 ; le système à base dix, décimal, en comporte dix : 0, 1, 2, ..., 9 ; les systèmes à base supérieure à dix comportent, outre les dix chiffres 0, 1, ..., 9, d'autres symboles : on utilise des lettres grecques, α, β, ... ; par exemple, en base douze, α représente dix et β onze).

◆ Ne pas confondre nombre et chiffre : le chiffre n'est qu'un symbole servant à l'écriture du nombre.

Quand, dans le système décimal, on écrit 1 682, cela veut dire que l'on a fait la somme de : 2 unités, 8 dizaines, 6 centaines et 1 millier, autrement dit, que :

$$1\,682 = 1 \times 10^3 + 6 \times 10^2 + 8 \times 10 + 2.$$

Ce qui est vrai pour dix est vrai pour tout entier naturel a supérieur ou égal à deux. On démontre que, si a est un entier naturel supérieur ou égal à deux, pour tout entier naturel x on peut trouver un entier naturel n (unique) et $n + 1$ entiers $x_n, x_{n-1}, \ldots, x_1, x_0$ compris entre 0 et a (déterminés de façon unique) tels que x puisse s'écrire :

$$x = x_n . a^n + x_{n-1} . a^{n-1} + \ldots + x_2 . a^2 + x_1 . a + x_0.$$

On convient alors d'écrire $x = x_n\, x_{n-1} \ldots x_1\, x_0$ quand il n'y a aucune confusion à craindre ; sinon, on indique en indice dans quel système de numération on s'est placé : $(121)_4$ (l'indice 4 indique qu'on est dans le système de numération de base quatre).

Dans n'importe quel système de numération à base a, un nombre inférieur à a s'écrit avec un seul chiffre ; a s'écrit 10 et une puissance n de a s'écrit 10000 ... 0 (1 suivi de n zéros).

Deux problèmes se présentent : comment traduire dans un système de base a quelconque (autre que dix) un nombre écrit dans le système décimal ? Comment transcrire dans le système décimal un nombre écrit dans un système de base a ? Supposons vouloir écrire 1245 dans le système à base a : on divise 1245 par a, puis le quotient obtenu par a, ..., jusqu'à ce qu'on arrive à un quotient (non nul) inférieur à a : ce dernier quotient sera le premier chiffre (à gauche) et les autres chiffres seront les restes trouvés successivement au cours des divisions effectuées.

Exemple : si on veut écrire 1245 en base deux. Disposition pratique du calcul :

1245	2										
1	622	2									
	0	311	2								
		1	155	2							
			1	77	2						
				1	38	2					
					0	19	2				
						1	9	2			
							1	4	2		
								0	2	2	
									0	1	2
										1	0

$(1245)_{10} = (10011011101)_2$

1245 s'écrit 10011011101 en base deux (on constate que quatre chiffres seulement étaient nécessaires pour l'écrire en base dix, alors qu'il en faut onze en base deux).

Supposons maintenant connaître le nombre 11001101 dans la base deux et vouloir le transcrire en base dix. Il suffit de se souvenir que 11001101 est l'écriture condensée de :

$$1 \times 2^7 + 1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2 + 1.$$

Il n'y a plus qu'à remplacer 2^7, 2^6, ..., par leur valeur en base dix :

$$(11001101)_2 = 128 + 64 + 8 + 4 + 1 = 205.$$

Pour traduire dans une base b un nombre écrit dans une base a, on pourrait procéder directement, mais notre grande habitude du système décimal nous permet d'aller plus vite en passant par l'intermédiaire du système décimal : le nombre écrit en base a est traduit dans le système décimal, puis retraduit en base b. Si on veut écrire en base deux le nombre écrit 356 en base huit, on commence par écrire $(356)_8$ dans le système décimal :

$(356)_8 = 3 \times 8^2 + 5 \times 8 + 6 = 238,$

puis on écrit 238 dans le système binaire :

238	2
0	119
	1

119	2
1	59
	1

59	2
1	29
	1

29	2
1	14
	0

14	2
0	7
	1

7	2
1	3
	1

3	2
1	1
	1

1	2
1	0

$(238)_8 = (11101110)_2$

Le grand intérêt présenté par les systèmes de numération de base a est de permettre d'effectuer facilement les opérations : addition, soustraction, sur les représentations chiffrées des nombres (ce que ne permet pas le système de numération romain, par exemple).

Opérations dans un système de numération à base quelconque

Les opérations se font strictement de la même façon que dans le système décimal (puisqu'en fait le système décimal n'est qu'un cas particulier de tous ces systèmes), en n'oubliant pas de faire jouer le rôle de la dizaine à la base a.

Voici quelques exemples d'opérations effectuées dans le système binaire ◆ (dans la multiplication et les divisions des exemples figurent les opérations intermédiaires, mais la disposition normale ne les comporte pas).

◆ Dans le système binaire : $1 + 1 = 10$.

1010	1110	1101	1101
+ 101	+ 111	− 100	− 111
1111	10101	1001	110

10111	111011	101
× 101	−101000	1011
10111	10011	
10111.	− 1010	
1110011	1001	
	− 101	
	100	

111011	101
1001	1011
1001	
100	

Ecriture d'un nombre rationnel dans un système de base quelconque

Supposons vouloir écrire $\frac{5}{8}$ en base deux :

$5 = 1 \times 2^2 + 1$ s'écrit 101
$8 = 2^3$ s'écrit 1 000

} on convient d'écrire $\frac{101}{1000} = 0{,}101.$

Si on veut écrire $\frac{3}{5}$ en base deux :

$3 = 1 \times 2 + 1$ s'écrit 11
$5 = 1 \times 2^2 + 1$ s'écrit 101
} $\frac{3}{5}$ s'écrit 0,10011001100...

On obtient simplement ces développements en effectuant les divisions du numérateur par le dénominateur (écrits dans le système considéré).
Dans le système de numération de base a : les nombres rationnels dont le dénominateur est une puissance de a ont un nombre fini de chiffres après la virgule (comme les nombres décimaux dans le système décimal), les autres en ont un nombre infini.

OCTAÈDRE : voir Polyèdres

OCTOGONE : voir Polygones

OPÉRATEUR : voir Loi de composition externe

OPPOSÉ : voir Symétrique

ORBITE page 45.

ORDONNÉ : voir aussi Ordre.

Ensemble ordonné
Un ensemble E muni d'une relation d'ordre $\mathcal{R}$ est dit ordonné par $\mathcal{R}$. Si la relation $\mathcal{R}$ est une relation d'ordre total, E est dit totalement ordonné ; si $\mathcal{R}$ est une relation d'ordre partiel, E est dit partiellement ordonné ; si $\mathcal{R}$ est une relation de bon ordre, E est dit bien ordonné.

Groupe ordonné
Soit E un ensemble muni d'une loi de composition interne notée $\divideontimes$ qui en fait un groupe. On dit que E est un groupe ordonné s'il est muni d'une relation d'ordre notée $\leqslant$ compatible avec la loi $\divideontimes$, c'est-à-dire que, pour tout x de E, si $a \leqslant b$, alors :

$$a \divideontimes x \leqslant b \divideontimes x \quad \text{et} \quad x \divideontimes a \leqslant x \divideontimes b.$$

Exemple : $\mathbb{Z}$ (ensemble des nombres entiers positifs et négatifs) est un groupe additif ordonné, car, si on a $a \leqslant b$, pour tout x de $\mathbb{Z}$:

$$a + x \leqslant b + x \quad \text{et} \quad x + a \leqslant x + b.$$

Anneau ordonné
Soit E un ensemble qui, muni des deux lois de composition interne notées $+$ et $\times$, a une structure d'anneau. On dit que E est un anneau ordonné s'il est muni d'une relation d'ordre notée $\leqslant$ possédant les propriétés suivantes :

si $a \leqslant b$, quel que soit x de A, $a + x \leqslant b + x$
($x + a < x + b$, puisque $a + x = x + a$ et $b + x = x + b$)
si $a \leqslant b$, quel que soit l'élément positif x de E,
$a \times x \leqslant b \times x$ et $x \times a \leqslant x \times b$.

Exemple : $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ sont des anneaux ordonnés ; il n'existe pas de relation d'ordre sur $\mathbb{C}$ qui en fasse un anneau ordonné.

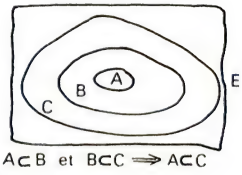
ORDONNÉE : voir Coordonnées

ORDRE pages 29, 32 à 34, 276, 282, 283, 309, 379, 394.

Dans l'ensemble $\mathcal{P}(E)$ des parties d'un ensemble E, la relation « A est inclus dans B » est telle que :

— quel que soit le sous-ensemble A de E, A est inclus dans lui-même :

$$A \subset A$$



— quels que soient les sous-ensembles A et B de E, si A est inclus dans B et si B est inclus dans A, alors A est égal à B :

$$A \subset B \text{ et } B \subset A \Longrightarrow A = B$$

— quels que soient les sous-ensembles A, B, C de E, si A est inclus dans B et B inclus dans C, alors A est inclus dans C :

$$A \subset B \text{ et } B \subset C \Longrightarrow A \subset C.$$

Une telle relation est une relation d'ordre.

D'une façon générale, une relation $\mathcal{R}$ définie sur un ensemble E est une relation d'ordre si elle est à la fois :

— réflexive : $\forall x \in E, \quad x \mathcal{R} x$

— antisymétrique : $\forall x \in E, \ \forall y \in E, \ x \mathcal{R} y$ et $y \mathcal{R} x \Longrightarrow x = y$

— et transitive : $\forall x \in E, \forall y \in E, \forall z \in E, \ x \mathcal{R} y$ et $y \mathcal{R} z \Longrightarrow x \mathcal{R} z$.

La relation d'égalité est aussi une relation d'ordre dans un ensemble E.
Dans l'ensemble des lettres de l'alphabet français, l'« ordre alphabétique » correspond à une relation d'ordre qui pourrait être définie par la phrase : « La lettre... est avant la lettre... dans l'ordre alphabétique. »

Ordre total

On dit qu'une relation d'ordre est une relation d'ordre total si, étant donné deux éléments quelconques x et y de E, on peut affirmer soit que x est en relation avec y, soit que y est en relation avec x ; c'est le cas de l'ordre alphabétique ; étant donné deux lettres, il y en a toujours une des deux qui est avant l'autre dans l'alphabet. On note souvent une relation d'ordre total $\leqslant$ (lire : inférieur ou égal à). A partir d'une telle relation, on peut en définir une autre, que l'on note $<$ (lire : strictement inférieur à), le fait que l'on ait $x < y$ signifiant que x est en relation avec y mais ne peut être égal à y. Une telle relation n'est plus une relation d'ordre (puisqu'elle n'est pas réflexive), mais on l'appelle néanmoins relation d'ordre strict.

Ordre partiel

Par contre, dans $\mathcal{P}(E)$, ensemble des parties de E, il existe des couples (A, B) de sous-ensembles tels qu'on n'ait ni $A \subset B$ ni $B \subset A$. Par exemple :

$$E = \{a, b, c\} \qquad A = \{a\} \qquad B = \{b, c\}.$$

On dit qu'une telle relation est une relation d'ordre partiel, et il en est de même pour toute relation d'ordre pour laquelle il existe des couples d'éléments (x, y) dont on ne puisse affirmer ni $x \mathcal{R} y$ ni $y \mathcal{R} x$.

ORIENTATION, ORIENTÉ page 55.

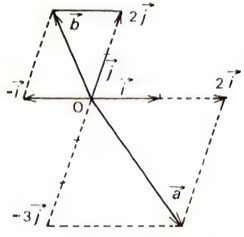
Soit E l'espace vectoriel, de dimension 2, des vecteurs du plan, d'origine O, dont une base est constituée par les 2 vecteurs notés $\vec{i}$ et $\vec{j}$. Si on considère une autre base de E, constituée des deux vecteurs $\vec{a} = 2\vec{i} - 3\vec{j}$ et $\vec{b} = -\vec{i} + 2\vec{j}$, le déterminant des composantes des vecteurs $\vec{a}$ et $\vec{b}$ dans la base $(\vec{i}, \vec{j})$ est :

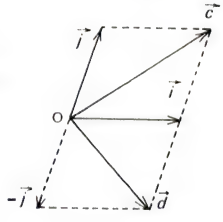
$$\begin{vmatrix} 2 & -1 \\ -3 & 2 \end{vmatrix} = 2 \times 2 - (-1)(-3) = 1.$$

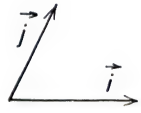
Il est positif ; on dit que les deux bases ont la même orientation.
Par contre, si on considère encore une autre base de E, celle qui est constituée par les deux vecteurs $\vec{c}$ et $\vec{d}$, avec $\vec{c} = \vec{i} + \vec{j}$ et $\vec{d} = \vec{i} - \vec{j}$, le déterminant des composantes de $\vec{c}$ et $\vec{d}$ dans la base $(\vec{i}, \vec{j})$ est :

$$\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = 1 \times (-1) - 1 \times 1 = -2.$$

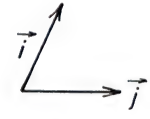
Il est négatif ; on dit que les deux bases sont d'orientations différentes.
Il y a deux orientations possibles de l'espace vectoriel des vecteurs du plan, d'origine O. On dit qu'on a orienté cet espace vectoriel si l'on a choisi l'une de ces deux orientations pour orientation privilégiée : on appellera bases de sens direct toutes les bases ayant cette orientation. Les bases de l'autre sont alors qualifiées de bases de sens rétrograde.







base $(\vec{i}, \vec{j})$
de sens direct



base $(\vec{i}, \vec{j})$
de sens rétrograde

ORTHOCENTRE : voir Hauteur

ORTHOGONAL (Repère) : voir Repère

Base orthogonale : voir Produit scalaire.

ORTHONORMÉ (Repère) : voir Repère

Base orthonormée : voir Produit scalaire.

OU (Disjonction) pages 98, 265, 310, 364, 415.

Si p et q sont deux propositions, la proposition « p ou q » est la proposition qui est vraie si et seulement si l'une au moins des deux propositions p ou q l'est.
Cette proposition est aussi appelée disjonction des deux propositions p et q ; parfois, on précise : disjonction inclusive. Sa table de vérité est représentée ci-contre.

p	q	p ou q
V	V	V
V	F	V
F	V	V
F	F	F

On peut aussi définir à partir des deux propositions p et q la proposition « p ou bien q » qui est vraie quand une seule des deux propositions est vraie. On appelle cette proposition disjonction exclusive. Sa table de vérité est représentée ci-contre.

p	q	p ou bien q
V	V	F
V	F	V
F	V	V
F	F	F

OUVERT : voir Topologie et Intervalle et la page 306.

P

PAIRE page 269.

Une paire est un ensemble formé de deux éléments. Il ne faut pas confondre paire avec couple ◆ ; il y a une seule paire $\{ x, y \}$, mais deux couples formés des mêmes éléments (distincts) x et y : le couple (x, y) et le couple (y, x).

◆ Voir aussi l'explication de cette différence dans l'article « les Ensembles », page 268.

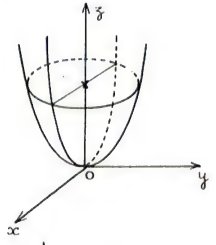
PARABOLE : voir Coniques et la page 225.

PARABOLOÏDE

Surface dont l'équation est (dans des axes de coordonnées convenablement choisis) de la forme :

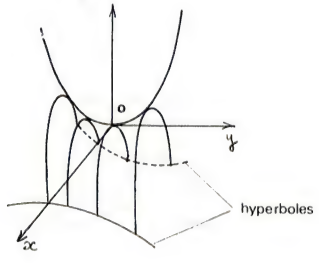
— soit $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2pz = 0$ (c'est alors un paraboloïde elliptique) ;

— soit $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 2pz = 0$ (c'est alors un paraboloïde hyperbolique).



Paraboloïde elliptique

Les sections par des plans parallèles à xOy sont des ellipses, et les sections par des plans parallèles à xOz ou à zOy sont des paraboles. Si les nombres a et b sont égaux, le paraboloïde est de révolution. Les sections planes parallèles à xOy sont des cercles. Un tel paraboloïde peut être engendré par la rotation d'une parabole autour de son axe.



Paraboloïde hyperbolique

Les sections par des plans parallèles à xOy sont des hyperboles, et les sections par des plans parallèles à xOz et à yOz sont des paraboles.

PARALLÉLÉPIPÈDE : voir Polyèdres

PARALLÉLOGRAMME : voir Polygones

PARAMÈTRE

Nombre qui figure dans une expression mathématique et dont on ne précise pas la valeur.

Paramètre d'une parabole : voir Coniques (paraboles).

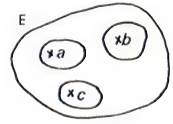
PARAMÉTRIQUE : voir Equations d'une courbe et d'une surface

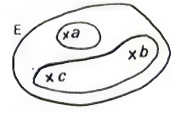
PARTIE ENTIÈRE : voir Fonction et la page 312.

PARTIEL (Ordre) : voir Ordre pages 282, 415.

PARTITION

Deux partitions possibles de E :





Une partition de l'ensemble E est une famille de parties non vides de E, disjointes deux à deux et dont la réunion est égale à E.

Par exemple, si $E = \{a, b, c\}$, la famille constituée par les trois parties $\{a\}$, $\{b\}$, $\{c\}$ est une partition de E, car :

$$\begin{cases} \{a\} \cap \{b\} = \{b\} \cap \{c\} = \{a\} \cap \{c\} = \varnothing \\ \{a\} \cup \{b\} \cup \{c\} = E. \end{cases}$$

De même, les deux parties $\{a\}$ et $\{b, c\}$ constituent une partition de E. (De façon générale, si A est une partie de E non vide et non égale à E, A et son complémentaire constituent une partition de E ◆.)

Lorsque $A_1, A_2, \ldots, A_n$ est une partition de E, tout élément de E appartient à l'un des sous-ensembles $A_1, A_2, \ldots, A_n$ et à un seul d'entre eux et, réciproquement, si une famille de sous-ensembles non vides de E est telle que tout élément de E appartienne à un et un seul d'entre eux, alors cette famille est une partition de E.

◆ A et son complémentaire $\complement_E A$ constituent une partition de E :

PASSAGE (Matrice de) ◆

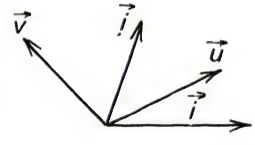
◆ Lire d'abord l'article « l'Algèbre linéaire », pages 55 à 57.

Considérons, dans l'espace vectoriel des vecteurs du plan, d'origine O, les deux bases B (constituée des deux vecteurs $\vec{i}$ et $\vec{j}$) et B′ (constituée des deux vecteurs $\vec{u}$ et $\vec{v}$). Un vecteur $\vec{X}$ de cet espace vectoriel a pour composantes (x, y) dans la base B, (x', y') dans la base B′ ; il est intéressant de savoir quelles sont les relations qui lient les nombres x et y aux nombres x' et y'.

Supposons que le vecteur $\vec{u}$ s'écrive : $\vec{u} = a\vec{i} + b\vec{j}$

et que le vecteur $\vec{v}$ s'écrive : $\vec{v} = a'\vec{i} + b'\vec{j}$

(autrement dit, les composantes de $\vec{u}$ dans la base B sont a et b, celles de $\vec{v}$ dans la base B sont a' et b').



Le vecteur $\vec{X}$ qui a pour composantes x' et y', dans la base B', s'écrit donc :

$$\vec{X} = x'\vec{u} + y'\vec{v},$$

c'est-à-dire :

$$\vec{X} = x'(a\vec{i} + b\vec{j}) + y'(a'\vec{i} + b'\vec{j}) = (x'a + y'a')\vec{i} + (x'b + y'b')\vec{j}.$$

Or $\vec{X}$ s'écrit aussi : $\vec{X} = x\vec{i} + y\vec{j}$

(puisque ses composantes dans la base B sont x et y). Les composantes d'un vecteur sur une base étant déterminées de façon unique◆, on a nécessairement :

◆ Voir l'article « l'Algèbre linéaire », page 69.

$$\begin{cases} x = x'a + y'a' \\ y = x'b + y'b' \end{cases}$$

ou, en utilisant l'écriture matricielle : $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & a' \\ b & b' \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$.

La matrice colonne des composantes du vecteur $\vec{X}$ dans la base B est égale au produit de la matrice colonne des composantes du vecteur $\vec{X}$ dans la base B' par la matrice P dont les éléments des colonnes sont les composantes des vecteurs de B' dans la base B. Cette matrice P est appelée matrice de passage de la base B à la base B'. On obtient ainsi les composantes de $\vec{X}$ dans la base B en fonction des composantes de $\vec{X}$ dans la base B'.
La matrice de passage de la base B' à la base B est la matrice inverse de cette matrice P et elle permet de calculer les composantes de $\vec{X}$ dans la base B' en fonction de celles de $\vec{X}$ dans la base B.

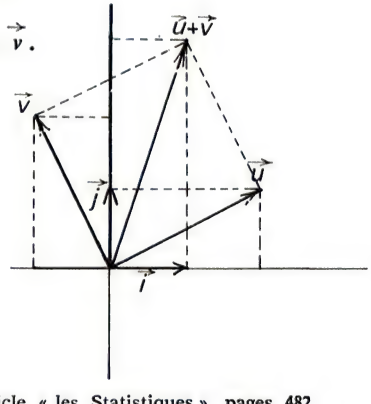
Exemple

$$\begin{cases} \vec{u} = 2\vec{i} + \vec{j} \\ \vec{v} = -\vec{i} + 2\vec{j} \end{cases} \qquad P = \begin{pmatrix} \underbrace{2}_{\alpha} & \underbrace{-1}_{\beta} \\ 1 & 2 \end{pmatrix}$$

α = composantes de $\vec{u}$; β = composantes de $\vec{v}$.

Le vecteur $\vec{X} = \vec{u} + \vec{v}$ qui a pour composantes (1, 1) dans la base $(\vec{u}, \vec{v})$ a pour composantes dans la base $(\vec{i}, \vec{j})$: x et y, donnés par :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \times 1 + (-1) \times 1 \\ 1 \times 1 + 2 \times 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$



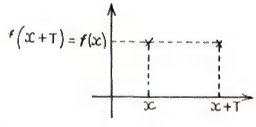
PEARSON (Corrélation de) : voir l'article « les Statistiques », pages 482, 483, 505 à 507.

PENTAGONE : voir Polygones

PENTE (d'une droite) : voir Coefficient directeur et la page 126.

PÉRIODE

Si f est une fonction réelle de la variable réelle x, on dit que f est périodique de période T (T étant un nombre réel) si la valeur de la fonction est la même pour x et $x + T$ (quel que soit le nombre x réel) : $f(x + T) = f(x)$.
Si T est une période, tout nombre réel égal au produit de T par un nombre entier est aussi une période.
Si la fonction f n'est pas constante et si elle admet des périodes non nulles, l'ensemble de toutes les périodes positives de cette fonction admet un plus petit élément ◆ : c'est celui-ci qu'on appelle en général période de la fonction.
Ainsi, les fonctions sinus cosinus sont des périodiques de période 2π ; la fonction tangente est périodique de période π.



◆ Toutes les périodes de la fonction sont des multiples de ce nombre.

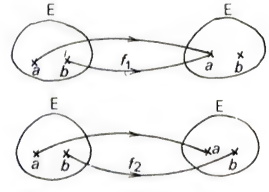
PERMUTATION◆ page 327.

◆ Lire d'abord Application, Arrangement, Image.

Si on cherche toutes les applications de $E = \{a, b\}$ dans lui-même, on en trouve quatre (2^2) :

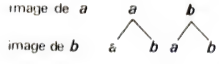
$f_1 : a \longmapsto a$ et $b \longmapsto a$ $\quad f_3 : a \longmapsto b$ et $b \longmapsto a$

$f_2 : a \longmapsto a$ et $b \longmapsto b$ $\quad f_4 : a \longmapsto b$ et $b \longmapsto b$



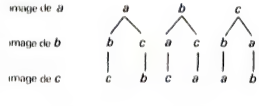
Les deux applications f_1 et f_4 ne sont pas injectives ◆, puisque les deux éléments a et b ont la même image (a pour f_1 et b pour f_2) ; par contre, f_2 et f_3 sont injectives (et même bijectives). Les deux applications f_2 et f_4 sont appelées permutations de E.
De la même façon, on appelle permutation d'un ensemble E (fini à n éléments) toute bijection de E sur lui-même. Toute injection de E sur lui-même étant une bijection (et toute bijection étant, d'abord, une injection), le nombre des bijections de E sur E est égal au nombre des injections de E sur E, c'est-à-dire au nombre d'arrangement des n éléments de E pris n à n. Ce nombre, noté P_n, est donc égal à $P_n = n(n-1)(n-2) \ldots (n-n+1) = n!$ (factorielle n). Dans le cas d'un ensemble à trois éléments, il y a $3! = 1.2.3 = 6$ permutations :

$f_1 \begin{cases} a \longmapsto a \\ b \longmapsto b \\ c \longmapsto c \end{cases}$ $f_2 \begin{cases} a \longmapsto a \\ b \longmapsto c \\ c \longmapsto b \end{cases}$ $f_3 \begin{cases} a \longmapsto b \\ b \longmapsto a \\ c \longmapsto c \end{cases}$



◆ Voir Injective.

$f_4 \begin{cases} a \longmapsto b \\ b \longmapsto c \\ c \longmapsto a \end{cases}$ $f_5 \begin{cases} a \longmapsto c \\ b \longmapsto b \\ c \longmapsto a \end{cases}$ $f_6 \begin{cases} a \longmapsto c \\ b \longmapsto a \\ c \longmapsto b \end{cases}$



Il arrive que l'on confonde l'application avec le résultat et que l'on nomme permutations de E non les applications $f_1, f_2, \ldots$, mais les résultats *abc*, *acb*, *bac*, *bca*, ...

P.G.C.D. (Plus grand commun diviseur) page 98.

P.G.C.D. de deux nombres entiers naturels (non nuls)

Si a et b sont deux nombres entiers naturels qui ne sont pas premiers ◆ entre eux, l'ensemble de leurs diviseurs communs comporte d'autres éléments que le nombre 1. Cet ensemble admet un plus grand élément qui est appelé plus grand diviseur commun de a et b ou P.G.C.D de a et b. Par exemple, 12 et 16 ont pour diviseurs respectivement 1, 2, 3, 4, 6, 12 et 1, 2, 4, 8, 16 ; leurs diviseurs communs sont donc 1, 2 et 4. Leur P.G.C.D. est 4 ◆.

◆ Voir ce mot.

Si a et b sont premiers entre eux, leur seul diviseur commun est 1 ; on dit que leur P.G.C.D. est 1. Par exemple, 3 et 8 sont premiers entre eux, leur P.G.C.D. est 1.

◆ Les diviseurs communs à a et b sont tous diviseurs du P.G.C.D. de a et b. Ici : 1, 2 et 4 sont tous diviseurs de 4.

Recherche du P.G.C.D.

Il y a deux façons de procéder pour rechercher le P.G.C.D. de deux nombres a et b :

— par l'algorithme d'Euclide ◆ ;

◆ Voir Algorithme.

— par l'utilisation de la décomposition en produit de facteurs premiers.

Exemple : recherche du P.G.C.D de 360 et de 108 :

— par l'algorithme d'Euclide :

$360 : 108 = 3$, reste 36 ;

$108 : 36 = 3$, reste 0 ;

le P.G.C.D. de 360 et 108 est donc 36 ;

— par la décomposition en produit de facteurs premiers ◆ :

$360 = 2^3 \times 3^2 \times 5$;

$108 = 2^2 \times 3^3$.

◆ Pour la décomposition en facteurs premiers, voir Premier.

On obtient le P.G.C.D. en faisant le produit de tous les facteurs qui figurent à la fois dans les deux décompositions, affectés de l'exposant le plus bas avec lequel ils figurent. Ici, 2 figure dans les deux décompositions, avec les exposants 2 et 3 : il figure donc dans le P.G.C.D. avec l'exposant 2 ; 3 figure avec les exposants 2 et 3 : il figure donc avec l'exposant 2 dans le P.G.C.D. Quant à 5, il ne figure que dans l'une des deux, donc ne figure pas dans le P.G.C.D. ; le P.G.C.D. est donc :

$2^2 \times 3^2 = 36$.

Propriétés du P.G.C.D.

— Le P.G.C.D. de a et de b est égal au P.G.C.D. de b et de a.

— Le P.G.C.D. de a et de a est égal à a.

— On peut définir le P.G.C.D. de trois nombres (ou plus) de la même façon (en faisant la liste des diviseurs communs aux trois nombres). Pour calculer le P.G.C.D. de trois nombres a, b, c, on peut soit calculer le P.G.C.D. de a et du P.G.C.D. de b et c, soit calculer le P.G.C.D. de c avec le P.G.C.D. de a et b.

— Le P.G.C.D. de a et de 1 est égal à 1.

Le P.G.C.D. de deux nombres entiers naturels étant encore un nombre entier naturel, la loi qui fait correspondre à deux entiers a et b leur P.G.C.D. est une loi de composition interne dans $\mathbb{N}^*$ (ensemble des entiers naturels strictement positifs). La première propriété traduit le fait que cette loi est commutative, la troisième qu'elle est associative, la deuxième que chaque élément est idempotent et la quatrième que 1 est élément absorbant.

P.G.C.D. de deux polynômes

On a défini sur l'ensemble des polynômes à coefficients réels une division (la division euclidienne) qui a les mêmes propriétés que la division à une unité près des nombres entiers. On peut donc définir exactement de la même façon le P.G.C.D. de deux polynômes ◆.

Par exemple, les deux polynômes :

$$P = x^3 + 2x^2 - x - 2 \quad \text{et} \quad Q = x^3 + x^2 - x - 1$$

sont tels que :

$$x^3 + x^2 - x - 1 : x^3 + 2x^2 - x - 2 = 1, \text{ reste } x^2 - 1$$

$$x^3 + 2x^2 - x - 2 : x^2 - 1 = x + 2, \text{ reste } 0.$$

Le P.G.C.D. de P et Q est $x^2 - 1$.

On aurait pu trouver ce P.G.C.D. en remarquant que :

$$P = (x^2 - 1)(x + 2) \quad \text{et} \quad Q = (x^2 - 1)(x + 1),$$

donc le P.G.C.D. est $x^2 - 1$.

◆ Il y a toutefois une différence : si un polynôme R est un diviseur du polynôme P, tout polynôme de la forme $k.R$ (où k est un nombre réel quelconque) est aussi un diviseur de P. Les diviseurs de P sont donc définis à une constante multiplicative près. Il en sera de même du P.G.C.D. Toutefois, ce polynôme est unique si on ne considère que des polynômes dont le coefficient du terme de plus haut degré est égal à 1.

PI (π) page 372.

Le nombre réel représenté par la lettre grecque π exprime le rapport entre la longueur d'un cercle et son diamètre. Ce nombre est un nombre transcendant, dont une valeur approchée est 3,141 59.

PLAN : voir Espace affine dans l'article « l'Algèbre linéaire », page 62.

PLAN VECTORIEL : voir l'article « l'Algèbre linéaire », page 71.

POISSON (Distribution de, loi de) : voir l'article « les Statistiques », pages 491 à 494.

POLAIRES : voir Coordonnées

POLYÈDRES page 326.

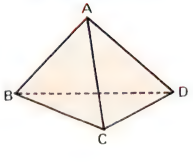
Étymologiquement, polyèdre signifie « plusieurs bases ». Le polyèdre est un volume limité par un nombre fini de plans. Les droites selon lesquelles se coupent ces plans sont appelées des arêtes. Les points d'intersection des arêtes sont des sommets. Si un des plans limitant le polyèdre contient deux arêtes qui ont un sommet commun, alors il contient un polygone dont les sommets sont des sommets du polyèdre. Ce polygone est appelé face du polyèdre (voir figures). Chaque plan contenant une face du polyèdre divise l'espace en deux régions ; si, pour chaque face, le polyèdre se trouve tout entier dans l'une des deux régions, on dit qu'il est convexe. Si le polyèdre est convexe, il existe entre le nombre a de ses arêtes, le nombre f de ses faces et le nombre s de ses sommets la relation d'Euler :

$$f + s = a + 2$$

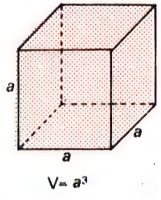
Les polyèdres réguliers sont ceux dont les faces sont des polygones réguliers. Il n'y en a que cinq :

— *Le tétraèdre régulier :* les quatre faces sont des triangles équilatéraux ; il a quatre sommets et quatre arêtes.

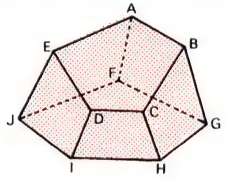
Tétraèdre régulier



— *Le cube :* les six faces sont des carrés ; il a huit sommets et douze arêtes. C'est un parallélépipède dont les six faces sont égales.
— *L'octaèdre régulier :* les huit faces sont des triangles équilatéraux, il a six sommets et douze arêtes.
— *Le dodécaèdre régulier :* les douze faces sont des pentagones réguliers, il a vingt sommets et trente arêtes.
— *L'isocaèdre :* les vingt faces sont des triangles équilatéraux, il a douze sommets et trente arêtes.

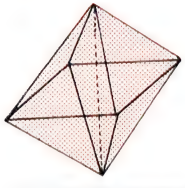


Cube

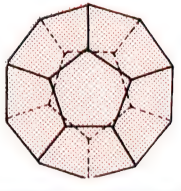


Heptaèdre

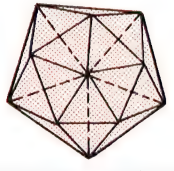
Octaèdre



Dodécaèdre

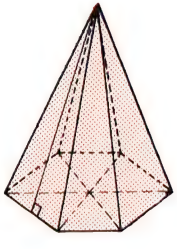


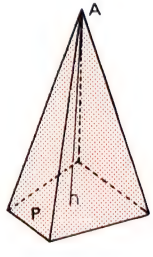
Icosaèdre



L'heptaèdre est un polyèdre à sept faces. Un polyèdre a au moins quatre faces. On obtient le nom du polyèdre en ajoutant la terminaison « -èdre » à un préfixe d'origine grecque indiquant le nombre de faces : tétraèdre (quatre faces), hexaèdre (6 faces ; parmi les hexaèdres figurent le cube et le parallélépipède).
Parmi les polyèdres, citons encore :

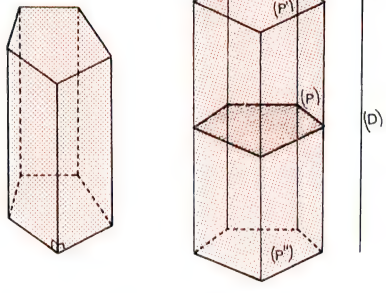
Pyramide





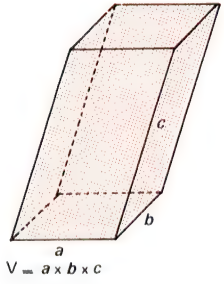
— *La pyramide :* c'est le volume obtenu à partir d'un polygone et d'un point A, le point A n'étant pas dans le plan du polygone P. Les faces sont des triangles de sommet A et dont la base est un côté du polygone P.
Si le polygone de base est régulier, la pyramide est régulière. Dans ce cas, les faces latérales sont des triangles isocèles. La hauteur issue du sommet est appelée apothème de la pyramide.

Prisme droit **Prisme**



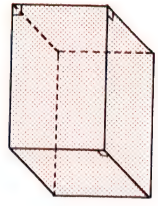
— *Le prisme.* Soient un polygone (P) et une droite (D) ; l'ensemble des droites parallèles à (D) et rencontrant (P) définit une surface (S). On appelle prisme le volume limité par (S) et deux plans parallèles au plan de (P) : (P') et (P'').
On appelle prisme droit un prisme dont les faces latérales sont des rectangles.

Parallélépipède



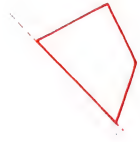
— *Le parallélépipède :* c'est un polyèdre dont les six faces sont deux à deux parallèles. Il a huit sommets et douze arêtes.
Un parallélépipède est rectangle si les faces sont des rectangles.

Parallélépipède rectangle

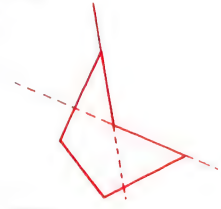


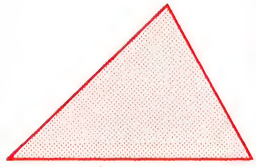
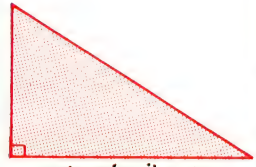
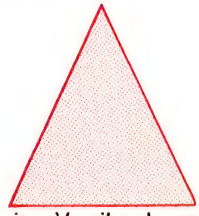
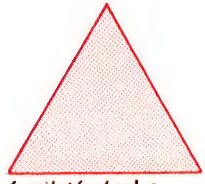
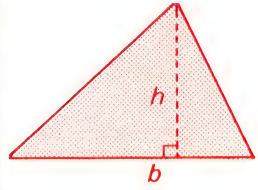
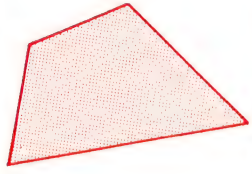
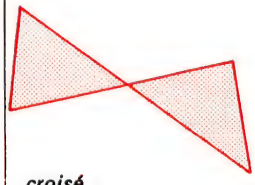
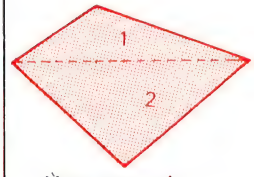
POLYGONES page 214.

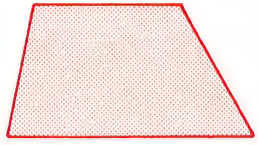
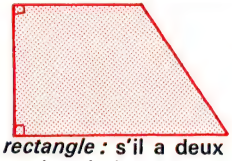
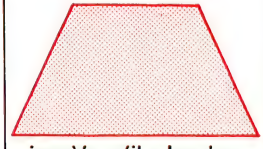
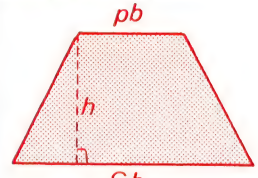
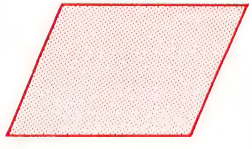
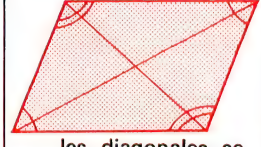
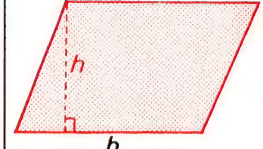
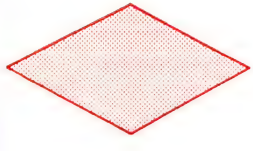
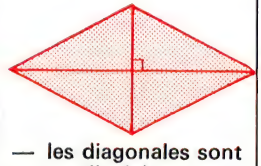
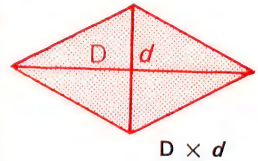
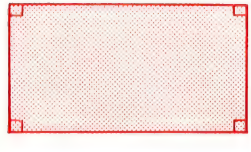
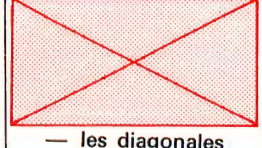
Étymologiquement : plusieurs angles.
Étant donné n points ($n \geqslant 3$) $A_1, A_2, \ldots, A_n$ situés dans un même plan et non alignés trois par trois, on appelle polygone de sommets $A_1, A_2, \ldots, A_n$ la figure formée par les n segments de droites $A_1A_2, A_2A_3, \ldots, A_{n-1}A_n, A_nA_1$. Les points $A_1, A_2, \ldots, A_n$ sont appelés les sommets du polygone et les segments $A_1A_2, A_2A_3, \ldots, A_nA_1$ les côtés du polygone.
Chaque côté du polygone délimite deux régions dans le plan. Si, pour chacun des côtés, le polygone se trouve tout entier dans la même région, il est qualifié de convexe ; il est qualifié de concave dans le cas contraire. On dit qu'il est croisé si deux côtés non consécutifs ont un point commun.
Si tous les côtés du polygone ont la même longueur, le polygone est dit régulier ; il est alors inscriptible.
Les polygones à 3 côtés sont appelés triangles. Les polygones à 4 côtés sont appelés quadrilatères. A partir de 5, on appelle pentagone, hexagone, ..., dodécagone (« gone » précédé d'un préfixe d'origine grecque indiquant le nombre) et, au-delà, il n'y a pas de nom particulier.

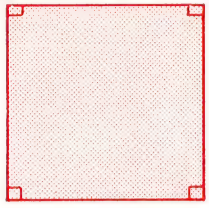
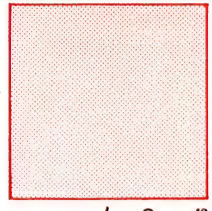
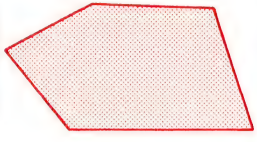
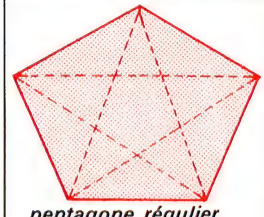
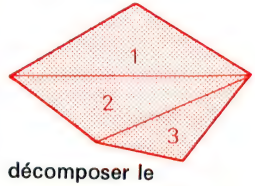
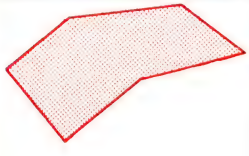
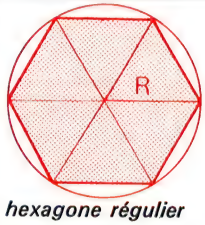
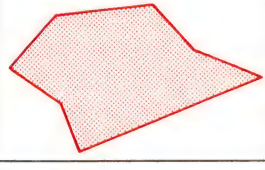
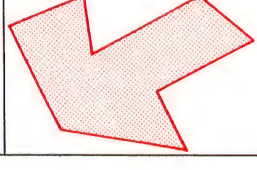

polygone convexe

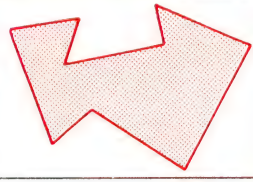
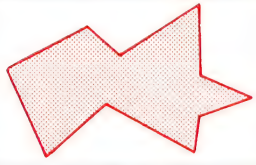
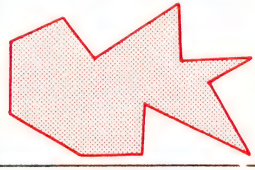
croisé


concave

nom	représentation	particularités	surfaces
Triangle trois sommets, trois côtés Voir aussi Hauteur, Hypoténuse, Bissectrice, Médiane, Médiatrice, Égalité	 *quelconque*	 ***rectangle :*** il a un angle droit  ***isocèle :*** il a deux côtés de même longueur  ***équilatéral :*** il a trois côtés de même longueur (le seul qui soit régulier)	 $S = \frac{1}{2} b \times h$
Quadrilatère quatre sommets, quatre côtés 1) **quadrilatère quelconque**		 *croisé*	 décomposer le quadrilatère en deux triangles : $S = S_1 + S_2$ Voir Triangle

nom	représentation	particularités	surfaces
2) **trapèze** s'il a deux côtés parallèles, appelés bases du trapèze		 *rectangle :* s'il a deux angles droits  *isocèle :* s'il a les deux côtés qui ne sont pas parallèles, de même longueur	 $S = \frac{(Gb + pb)}{2} \times h$
3) **parallélogramme** s'il a les quatre côtés deux à deux parallèles		 — les diagonales se coupent en leur milieu ; — les angles opposés sont égaux	 $S = b \times h$
4) **losange** parallélogramme dont les quatre côtés sont égaux		 — les diagonales sont perpendiculaires	Voir Parallélogramme ou :  $\frac{D \times d}{2}$
5) **rectangle** parallélogramme qui a un angle droit ⟹ les quatre angles sont droits		 — les diagonales sont égales	 $S = L \times l$

nom	représentation	particularités	surfaces
6) **carré** losange qui a un angle droit ⟹ les quatre angles sont droits, ou rectangle qui a deux côtés consécutifs de même longueur ⟹ les quatre côtés ont même longueur		— possède toutes les particularités du rectangle et du losange (voir ci-dessus)	
Pentagone cinq sommets, cinq côtés		 *pentagone régulier convexe* cinq côtés de même longueur	 décomposer le pentagone en polygones dont les surfaces sont connues : $S = S_1 + S_2 + S_3$
Hexagone six sommets, six côtés		 *hexagone régulier* six côtés de même longueur ; il est inscriptible dans un cercle de rayon égal au côté de l'hexagone	utiliser le même processus que celui qui est utilisé pour le pentagone (il en sera de même pour tous les autres polygones ayant un nombre de côtés supérieur à quatre)
Heptagone sept sommets, sept côtés			voir Pentagone
Octogone huit sommets, huit côtés			voir Pentagone

nom	représentation	particularités	surfaces
Ennéagone neuf sommets, neuf côtés			voir Pentagone
Décagone dix sommets, dix côtés			voir Pentagone
Dodécagone douze sommets, douze côtés			voir Pentagone

POLYGONE DE FREQUENCE : voir l'article « les Statistiques », page 462.

POLYNÔME pages 250, 251, 286 à 290, 290, 214, 221, 402, 411, 442.

Polynôme à une variable et à coefficients réels ◆

Le sens le plus simple (et le seul utilisé autrefois) du mot polynôme est « somme de monômes » (ainsi, si a, b, x sont des nombres réels, $ax^2 + b^2x + x^3$ est un polynôme). Si le polynôme est une somme de deux monômes, on l'appelle binôme ; s'il est une somme de trois monômes, on l'appelle trinôme.

Actuellement, on préfère distinguer :

— le polynôme P défini comme une suite de nombres réels, nuls à partir d'un certain rang $(a_0, a_1, a_2, \ldots, a_n, 0, 0, \ldots)$, qui sont appelés les coefficients du polynôme ;

— la fonction polynôme, qui à chaque nombre réel x fait correspondre le nombre réel $a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ (que l'on note $P(x)$).

Si le polynôme P est défini par la suite $(a_0, a_1, a_2, \ldots, a_n, 0, \ldots)$, où le dernier terme non nul est le n-ième, le nombre n est appelé le degré du polynôme. Dans le cas où tous les coefficients sont nuls, on dit que le polynôme P est le polynôme nul ; ce polynôme n'a pas de degré.

On peut définir sur l'ensemble des polynômes à coefficients réels plusieurs lois de composition (nous donnerons les formules avec les fonctions polynômes, plus faciles à utiliser).

◆ Nous nous limitons ici aux polynômes à coefficients réels. Néanmoins, toutes ces définitions sont encore valables si les coefficients et x appartiennent à un anneau commutatif unitaire A (dont les lois de composition interne sont notées + et ×), en particulier s'ils appartiennent au corps des nombres complexes $\mathbb{C}$.

Somme de deux polynômes

Si P et Q sont deux polynômes tels que :

$$P(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

$$Q(x) = b_0 + b_1x + b_2x^2 + \ldots + b_px^p,$$

la somme P + Q des deux polynômes est le polynôme tel que :

$$(P + Q)(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \dots$$

Exemple

$$P(x) = 1 - 2x + 3x^2 + x^3$$
$$Q(x) = 2 - 3x + 4x^2$$
$$(P + Q)(x) = 3 - 5x + 7x^2 + x^3.$$

Si le degré de P et le degré de Q sont différents, le degré du polynôme P + Q est égal au plus grand des deux degrés (de P et de Q) ; il en est ainsi de l'exemple ci-dessus. Si le degré de P et le degré de Q sont égaux, le degré de P + Q est inférieur ou égal au degré commun de P et de Q.

$P(x) = 1 - 2x + 3x^2$ | degré de P = degré de Q
$Q(x) = 2 + x - 3x^2$ | degré de P + Q < degré de P.
$(P + Q)(x) = 3 - x.$

$P(x) = 1 - 2x + 3x^2$ | degré de P = degré de Q
$Q(x) = 2 + x + x^2$ | = degré de P + Q.
$(P + Q)(x) = 3 - x + 4x^2.$

Produit de deux polynômes

Soient : $P(x) = 1 + 2x + x^2$

$Q(x) = 2 - 3x$

$$(P \times Q)(x) = (1 + 2x + x^2) \times (2 - 3x) \blacklozenge$$
$$= (1 \times 2) + (2x \times 2) + (x^2 \times 2) + 1 \times (-3x) + 2x \times (-3x) + x^2 \times (-3x)$$
$$= 2 + 4x + 2x^2 - 3x - 6x^2 - 3x^3$$
$$= 2 + x - 4x^2 - 3x^3.$$

◆ On développe ce produit en tenant compte de la distributivité de la multiplication par rapport à l'addition dans l'ensemble des nombres réels.

Si P est un polynôme tel que :

$$P(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

et si Q est un polynôme tel que :

$$Q(x) = b_0 + b_1x + b_2x^2 + \dots + b_px^p,$$

le produit des deux polynômes P et Q est le polynôme P × Q tel que :

$$(P \times Q)(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0)x^3 + \dots$$

Le degré du produit P × Q est égal à la somme du degré de P et du degré de Q. Ainsi, dans l'exemple ci-dessus :

degré de P = 2, degré de Q = 1, degré de P × Q = 3.

Muni de ces deux lois de composition interne ◆, l'ensemble des polynômes à coefficients réels est un anneau commutatif unitaire.

◆ Ces lois sont internes, puisque la somme (comme le produit) de deux polynômes à coefficients réels est encore un polynôme à coefficients réels.

Produit d'un polynôme par un nombre réel

Si P est un polynôme tel que :

$$P(x) = a_0 + a_1x + \dots + a_nx^n,$$

on appelle produit du polynôme P par le nombre réel k ◆ et on note kP le polynôme tel que :

$$kP(x) = k.a_0 + (k.a_1)x + (k.a_2)x^2 + \dots + (k.a_n)x^n.$$

◆ Cette loi est une loi de composition externe, puisque le nombre k est extérieur à l'ensemble des polynômes et que le résultat kP appartient à cet ensemble.

Muni de l'addition et du produit par un nombre réel, l'ensemble des polynômes à coefficients réels a structure d'espace vectoriel sur $\mathbb{R}$ (ensemble des nombres réels) ◆. Cet espace vectoriel n'est pas de dimension finie. Par contre, si on se limite aux polynômes de degré inférieur ou égal à n (n étant un nombre entier naturel fixé, différent de 0), l'addition est encore une loi de composition interne dans cet ensemble (la somme de deux polynômes de degré inférieur ou égal à n est encore un polynôme de degré inférieur ou égal à n) ; le produit par k réel d'un polynôme de degré inférieur ou égal à n est encore un polynôme de degré inférieur ou égal à n ; l'ensemble des polynômes de degré inférieur ou égal à n est un espace vectoriel sur $\mathbb{R}$, de dimension finie, égale à $n + 1$.

◆ Dans le cas où les coefficients appartiennent à un anneau A, la structure n'est plus celle d'espace vectoriel mais de module sur A.

Polynômes à plusieurs variables et à coefficients réels

Considérons par exemple les polynômes à deux variables x et y. On fait la même distinction entre le polynôme P et la fonction polynôme qui à tout couple (x, y) de nombres réels fait correspondre le nombre réel $P(x, y)$.
Si on considère ainsi le polynôme P tel que :

$$P(x, y) = 3x^2y + x^3y^2 - x^2y^4,$$

pour un tel polynôme, on distingue :

— le degré du polynôme par rapport à x, qui est égal au plus grand des exposants de x (ici, 3) ;

— le degré du polynôme par rapport à y, qui est égal au plus grand des exposants de y (ici, 4) ;

— le degré du polynôme par rapport à l'ensemble des deux variables x et y : la plus grande somme obtenue en ajoutant, pour chaque monôme, l'exposant de x et celui de y (ici, 6).

POPULATION :

voir l'article « les Statistiques », page 456.

POSITIF

Nombre positif : nombre supérieur à 0.
Soit E un ensemble ayant structure de groupe ◆ pour la loi de composition interne notée $\divideontimes$, dont l'élément neutre est noté 0 ; si ce groupe est un groupe ordonné pour la relation d'ordre notée $\leqslant$, tout élément x de E supérieur ou égal à 0 est positif, tout élément x de E strictement supérieur à 0 est strictement positif.

◆ Voir Groupe.

P.P.C.M. (Plus petit commun multiple)

P.P.C.M. de deux nombres entiers naturels

Si a et b sont deux nombres entiers naturels, l'ensemble de leurs multiples communs a un plus petit élément qui est appelé P.P.C.M. de a et de b.

Par exemple, 6 et 9 ont pour multiples :

6, 12, *18*, 24, 30, *36*, 42, 48, *54*, ...

et 9, *18*, 27, *36*, 45, *54*, 63, ...

Les multiples communs sont donc : 18, 36, 54, ... ◆. Le plus petit est 18.
Si a et b sont premiers ◆ entre eux, leur P.P.C.M. est égal au produit $a \times b$.

◆ Les multiples communs de deux nombres a et b sont tous les multiples du P.P.C.M. de ces deux nombres. Ici, ces multiples communs sont les multiples de 18.

◆ Voir ce mot.

Recherche du P.P.C.M.

Pour calculer le P.P.C.M. de a et de b, on utilise la décomposition en produit de facteurs premiers de a et de b ◆. Le P.P.C.M. est égal au produit de tous les facteurs figurant dans l'une ou l'autre des deux décompositions, chacun étant affecté de l'exposant le plus grand avec lequel il figure.

◆ Voir la décomposition en facteurs premiers à Premier.

Ainsi, le P.P.C.M. de 360 et de 108 est obtenu de la façon suivante :

$$360 = 2^3 \times 3^2 \times 5$$
$$108 = 2^2 \times 3^3.$$

Puisque 2 figure avec les exposants 2 et 3, il figure dans le P.P.C.M. avec l'exposant 3 ; puisque 3 figure avec les exposants 2 et 3, il figure dans le P.P.C.M. avec l'exposant 3 ; puisque 5 figure avec l'exposant 1, il figure dans le P.P.C.M. avec l'exposant 1.

Le P.P.C.M. est donc égal à :

$$2^3 \times 3^3 \times 5 = 1\,080.$$

Propriétés du P.P.C.M.

— Le P.P.C.M. de a et b est égal au P.P.C.M. de b et a.

— Le P.P.C.M. de a et a est égal à a.

— Le P.P.C.M. de a et 1 est égal à a.

— On peut définir le P.P.C.M. de plusieurs nombres de façon analogue, et, si on veut calculer le P.P.C.M. de trois nombres a, b, c, on peut soit calculer le P.P.C.M. de a et du P.P.C.M. de b et c, soit calculer le P.P.C.M. de c et du P.P.C.M. de a et b.

Le P.P.C.M. de deux entiers naturels étant encore un entier naturel, la loi qui associe à deux nombres a et b leur P.P.C.M. est une loi de composition interne dans $\mathbb{N}^*$ (ensemble des entiers naturels).

La première propriété indique que cette loi est commutative, la troisième que 1 est élément neutre, la quatrième que la loi est associative et la deuxième que tout élément est idempotent.

P.P.C.M. de deux polynômes

Comme on a défini la multiplication de deux polynômes ◆ et que cette multiplication a les mêmes propriétés que celles des nombres réels, on peut également parler du P.P.C.M. de deux polynômes. Ainsi, si :

◆ Voir Polynôme.

$$P = (x^2 - 1)(x + 2)$$

et $$Q = (x^2 - 1)(x + 1),$$

le P.P.C.M. de P et Q est le polynôme :

$$(x^2 - 1)(x + 2)(x + 1).$$

PREMIER (Nombre) pages 53, 326, 380.

Un nombre entier naturel n différent de 1 est un nombre premier s'il n'admet comme diviseurs que lui-même et 1.

Par exemple, 3 est un nombre premier, car il n'admet que 3 et 1 comme diviseurs ; 4 n'est pas premier, car il admet comme diviseurs 4, 2 et 1 (de même, aucun nombre pair n'est premier, car il admet toujours 2 comme diviseur).

Il existe une infinité de nombres premiers.

Il n'existe aucune loi permettant de dire, en voyant un nombre impair, s'il est premier ou non : il faut essayer de le diviser successivement par tous les nombres

premiers qui lui sont inférieurs en les essayant dans l'ordre croissant. En fait, il suffit d'essayer tous les nombres premiers dont le carré soit inférieur à ce nombre, mais c'est quand même fort long si le nombre est grand.

Exemple de recherche pour savoir si un nombre est premier ou pas. Nombre 113 ; essayons les diviseurs premiers possibles à partir de 3 :

113 n'est pas divisible par 3

113 n'est pas divisible par 5

113 n'est pas divisible par 7

113 n'est pas divisible par 11

(le carré de 11 est supérieur à 113, donc on s'arrête ici).

113 est un nombre premier.

Décomposition en facteurs premiers : tout entier naturel n peut être écrit, de façon unique, sous forme d'un produit de nombres premiers (si le nombre n est lui-même premier, le produit ne comporte qu'un terme : n). Par exemple, 12 se décompose de la manière suivante : $2 \times 2 \times 3$ (tous trois étant des nombres premiers) ; 7 se décompose en un produit qui ne comporte qu'un terme : 7.
Trouver ces nombres premiers s'appelle décomposer en produit de facteurs premiers.

Exemple de disposition pratique du calcul :

600	2
300	2
150	2
75	3
25	5
5	5
1	

$600 = 2^3 \times 3 \times 5^2$

Pour fabriquer une table donnant tous les nombres premiers de 1 à 100, par exemple :

~~0~~	~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~
~~10~~	11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19
~~20~~	~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29
~~30~~	31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~
~~40~~	41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	~~49~~
~~50~~	~~51~~	~~52~~	~~53~~	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59
~~60~~	61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~
~~70~~	71	~~72~~	73	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	79
~~80~~	~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89
~~90~~	~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~

On écrit tous les nombres de 1 à 100, puis on supprime tous les multiples de 2, puis ceux de 3 qui restent, puis ceux de 5 et, enfin, ceux de 7.

Nombres premiers de 1 à 100 :
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.

PRÉORDRE ◆ page 309.

◆ Il est conseillé de lire d'abord Relation, Réflexive et Transitive.

Une relation binaire sur un ensemble E est une relation de préordre si elle est réflexive et transitive.
Relations d'ordre et relations d'équivalence sont des relations de préordre.

PRIMITIVE (Fonction) ◆ pages 61, 143 à 147, 313, 349, 340.

◆ Voir aussi l'article « l'Analyse », page 143.

Soit f et F deux fonctions définies sur l'intervalle (a, b) de $\mathbb{R}$ (ensemble des nombres réels) et à valeurs réelles.
F est une primitive de f si :
— elle est dérivable en tout point x de (a, b) ;
— pour tout x de (a, b) la valeur de la dérivée de F, F′, est égale à f :

$$\forall\, x \in (a, b) \qquad F'(x) = f(x)$$

(lire : quel que soit x appartenant à l'intervalle (a, b), F′ de x est égale à $f(x)$).

Exemple ◆

◆ Voir le tableau des fonctions primitives à Fonction.

La fonction $F(x) = \frac{x^3}{3}$ définie sur $\mathbb{R}$ est une primitive de la fonction $f(x) = x^2$ définie sur $\mathbb{R}$, car la dérivée de $\frac{x^3}{3}$ est x^2 ; la dérivée de $\frac{x^3}{3} + 2$ est aussi x^2 ; donc, la fonction $G(x) = \frac{x^3}{3} + 2$ est aussi une primitive de $f(x) = x^2$. Il en est de même de toute fonction $\mathcal{F}$ de la forme :

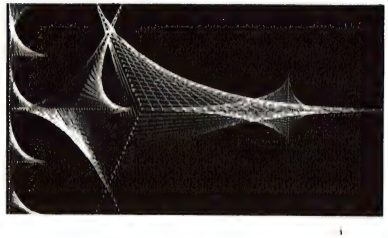
$$\mathcal{F}(x) = \frac{x^3}{3} + C,$$

où C est une constante quelconque.

De façon générale, si la fonction f admet une primitive F sur (a, b), elle en admet une infinité, chacune de ces primitives étant obtenue à partir de F par addition d'une constante.

PRISME : voir Polyèdres

PROBABILITÉS : voir l'article pages suivantes et les pages 163, 167 à 170, 171, 198, 456, 485.



Les probabilités

par Michèle Pascal

Le calcul des probabilités a pris naissance au XVII^{e} siècle à propos de l'étude des jeux de hasard. La correspondance de Pascal et de Fermat sur ce sujet est restée célèbre. Au XVIII^{e} siècle, de grands mathématiciens, tels Leibniz, de Moivre, Jacques Bernoulli, Laplace, Gauss ◆, se sont intéressés à cette branche des mathématiques et ont contribué à son développement. C'est au XIX^{e} siècle, avec Joseph Bertrand, Henri Poincaré, Émile Borel, que le calcul des probabilités a connu son plus grand développement.
Aujourd'hui, le calcul des probabilités est passé du domaine de la théorie dans celui des applications pratiques. En particulier, il est à la base de la méthode statistique qui s'introduit progressivement dans presque toutes les branches de la connaissance.

Michèle Pascal.
Professeur au lycée de Sèvres, membre diplômé de l'Institut des actuaires français.

◆ Voir l'article « les Statistiques », page 499.

LES ÉVÉNEMENTS

On appelle épreuve une expérience dont le résultat est aléatoire. Une épreuve peut donner lieu à un certain nombre d'événements dont l'un au moins se réalisera. Par exemple, jeter un dé est une épreuve qui peut engendrer plusieurs événements tels que : obtenir un as, un 6, un nombre pair, un nombre supérieur à 2.

Événements incompatibles

Deux événements seront appelés incompatibles s'ils ne peuvent se réaliser en même temps. Par exemple, lorsqu'on jette un dé : obtenir un 3 et obtenir un 4 sont deux événements incompatibles.

Événements élémentaires

Toutes les éventualités, tous les résultats possibles d'une épreuve s'appellent événements élémentaires. Par exemple :
— en jetant un dé, il y a six événements élémentaires qui sont ; obtenir 1, 2, 3, 4, 5, 6 ;

— en tirant une carte d'un jeu de 32 cartes, il y a 32 événements élémentaires possibles ; mais, si l'on s'intéresse seulement à la « couleur » de la carte, il n'y a que 4 événements élémentaires : tirer un cœur, un carreau, un pique, un trèfle.

Nous appellerons E l'ensemble des événements élémentaires relatifs à une épreuve ε :

$$E = \{ e_1, e_2, e_3, \ldots, e_n \}.$$

Si l'épreuve E consiste à jeter un dé en s'intéressant au numéro sorti, voici l'ensemble des événements élémentaires :

$$E = \{ 1, 2, 3, 4, 5, 6 \}.$$

Si l'épreuve ε' consiste à tirer une carte d'un jeu de 32 cartes en s'intéressant à la « couleur », on a :

$$E' = \{ \text{pique, cœur, carreau, trèfle} \}.$$

Les événements élémentaires déterminent une partition de E formée de sous-ensembles contenant chacun un élément de E et un seul. Leur intersection deux à deux est vide, ce qui revient à dire que *deux événements élémentaires sont incompatibles deux à deux :*

$$\{ e_1 \} \cap \{ e_2 \} = \varnothing \blacklozenge.$$

◆ $\varnothing$ = ensemble vide.

Les événements autres qu'élémentaires

A partir des événements élémentaires, on peut concevoir beaucoup d'autres événements. Tous ces événements sont des sous-ensembles de E. L'ensemble de tous ces événements est l'ensemble $\mathfrak{P}(E)$: ensemble des parties de E.

a) Union de deux ou plusieurs événements élémentaires ; on jette un dé et on envisage l'événement A : sortir un nombre supérieur à 3 :

$$A = \{ 4, 5, 6 \} \in \mathfrak{P}(E), \qquad \in : \text{appartient à}$$

ou l'événement B : sortir un multiple de 3 :

$$B = \{ 3, 6 \} \in \mathfrak{P}(E).$$

b) Union de deux ou plusieurs événements élémentaires ou non ; ces derniers peuvent être assemblés de deux façons :

— Par la conjonction « ou » : on jette un dé et l'on veut obtenir un nombre supérieur à 3 *ou* un nombre multiple de 3. L'événement souhaité : C, est réalisé dès que l'un au moins des deux événements A ou B est réalisé. Cela correspond à l'opération « union » dans $\mathfrak{P}(E)$ ◆ :

◆ $\cup$ = union.

$$C = A \cup B = \{ 4, 5, 6 \} \cup \{ 3, 6 \} = \{ 3, 4, 5, 6 \}.$$

L'union étant une opération interne dans $\mathfrak{P}(E)$, le résultat est toujours dans $\mathfrak{P}(E)$.
— Par la conjonction « et » : on jette un dé et l'on veut obtenir un nombre supérieur à 3 *et* multiple de 3. Dans ce cas, l'événement souhaité : D, est réalisé si et seulement si les deux événements le sont. Cela correspond à l'opération « intersection » dans $\mathfrak{P}(E)$◆ : ◆ ∩ = intersection.

$$D = A \cap B = \{ 4, 5, 6 \} \cap \{ 3, 6 \} = \{ 6 \}.$$

L'intersection étant une opération interne dans $\mathfrak{P}(E)$, le résultat est toujours dans $\mathfrak{P}(E)$.

L'événement impossible

Si nous assemblons deux événements incompatibles par la conjonction *et*, nous obtenons l'événement impossible, soit l'ensemble vide qui appartient à $\mathfrak{P}(E)$. Par exemple, A est l'événement : sortir un nombre supérieur à 3, B est l'événement : sortir un as. L'événement C qui consiste à sortir un nombre supérieur à 3 *et* un as est impossible :

$$C = A \cap B = \{ 4, 5, 6 \} \cap \{ 1 \} = \varnothing.$$

L'événement certain

Si nous considérons l'événement résultant de l'union de tous les événements élémentaires, nous obtenons l'événement certain. En jetant un dé, l'événement E, consistant à sortir les nombres 1 ou 2 ou 3 ou 4 ou 5 ou 6, est certain :

$$\begin{aligned} E &= \{ 1 \} \cup \{ 2 \} \cup \{ 3 \} \cup \{ 4 \} \cup \{ 5 \} \cup \{ 6 \} \\ &= \{ 1, 2, 3, 4, 5, 6 \}. \end{aligned}$$

Cet ensemble E appartient à $\mathfrak{P}(E)$.
Remarquons que l'événement certain est obtenu chaque fois que l'on unit par la conjonction *ou* des événements formant une partition de E. En particulier, l'union de deux événements contraires est une partition de E.

L'événement contraire

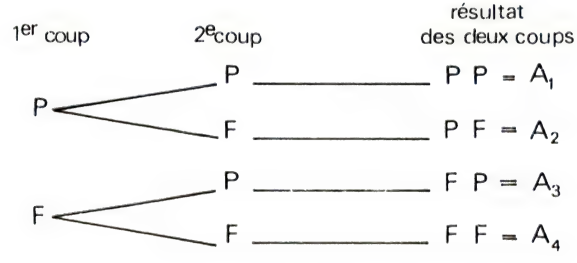
Deux événements d'une même épreuve sont des événements contraires s'ils sont incompatibles et si l'un d'eux se réalise certainement au cours de l'expérience. En jetant un dé : sortir un nombre inférieur à 6 est l'événement contraire de sortir un 6. Dans $\mathfrak{P}(E)$, un événement et son contraire correspondent à un élément de $\mathfrak{P}(E)$ et son complémentaire dans $\mathfrak{P}(E)$. Soient A et $\overline{A}$ deux événements contraires :

$$A \cup \overline{A} = E \qquad A \cap \overline{A} = \varnothing$$

Enfin, nous dirons que l'événement A implique l'événement B si, lorsque A est réalisé, B l'est aussi. Dans ce cas, A et B étant deux éléments de $\mathfrak{T}(E)$: $A \subset B$. En jetant un dé, l'événement A : sortir un 2, implique l'événement B : sortir un nombre pair :

$$\{ 2 \} \subset \{ 2, 4, 6 \}.$$

L'ensemble $\mathfrak{T}(E)$ muni de la réunion et de l'intersection constitue un modèle d'algèbre de Boole et, par suite, l'ensemble des événements relatifs à une épreuve ε dont le nombre des événements élémentaires est fini constitue également un modèle d'algèbre de Boole. Cette remarque permet d'utiliser les propriétés de l'algèbre de Boole, sans tenir compte des particularités de chaque épreuve. Il suffira d'établir la liste des événements élémentaires.
L'exemple suivant est destiné à montrer comment un événement quelconque relatif à une épreuve déterminée appartient à $\mathfrak{T}(E)$. Soit une épreuve consistant à jeter deux fois une pièce de monnaie. L'ensemble des événements élémentaires peut s'établir en faisant un « arbre » :



$$E = \{ PP, PF, FP, FF \} = \{ A_1, A_2, A_3, A_4 \}.$$

$$\mathfrak{T}(E) = \{ \varnothing, \{ A_1 \}, \{ A_2 \}, \{ A_3 \}, \{ A_4 \}, \{ A_1, A_2 \}, \{ A_1, A_3 \}, \{ A_1, A_4 \}, \{ A_2, A_3 \}, \{ A_2, A_4 \}, \{ A_3, A_4 \}, \{ A_1, A_2, A_3 \}, \{ A_1, A_2, A_4 \}, \{ A_2, A_3, A_4 \}, \{ A_1, A_3, A_4 \}, \{ A_1, A_2, A_3 A_4 \} \}.$$

Envisageons maintenant quelques événements de ce jeu exprimés dans le langage des ensembles :

— Sortir 2 piles :

$\{ A_1 \} \in \mathfrak{T}(E)$ lire : l'ensemble A_1 appartient aux parties de E.

— Sortir 2 piles ou 2 faces :

$$\{ A_1 \} \cup \{ A_4 \} = \{ A_1, A_4 \} \in \mathfrak{T}(E).$$

— Sortir face au premier coup :

$$\{ A_3 \} \cup \{ A_4 \} = \{ A_3, A_4 \} \in \mathfrak{T}(E).$$

— Sortir au moins une pile :

$$\{ A_2 \} \cup \{ A_3 \} \cup \{ A_1 \} = \{ A_1, A_2, A_3 \} \in \mathfrak{T}(E).$$

— Sortir 2 faces (c'est l'événement contraire du précédent) :

$$\{ \overline{A_1, A_2, A_3} \} = \{ A_4 \} \in \mathfrak{T}(E).$$

— Sortir 2 piles et 2 faces (événement impossible) :

$$\{ A_1 \} \cap \{ A_4 \} = \varnothing \in \mathfrak{T}(E).$$

Le nombre d'événements élémentaires est infini

Il est parfois impossible de dénombrer les événements élémentaires d'une épreuve. Soit, par exemple : on jette une pièce de monnaie sur une table ; l'épreuve consiste à faire coïncider le centre de la pièce avec un point de la table. Il y a une infinité d'événements élémentaires possibles, puisqu'il y a une infinité de points. Cette épreuve n'est pas concernée par tout ce qui précède, qui ne s'adresse qu'aux ensembles finis ou dénombrables. Il est cependant possible d'étendre aux ensembles infinis d'événements la notion d'algèbre de Boole. Nous ne l'envisagerons pas ici. Signalons seulement qu'une épreuve de ce genre relèvera de la théorie des probabilités continues, alors que toutes les autres épreuves, dont le nombre des événements élémentaires est soit fini, soit dénombrable, relèvent des probabilités discrètes.

LA NOTION DE PROBABILITÉ

La notion de probabilité est le résultat d'un raisonnement dans lequel on évalue le nombre de chances d'obtenir la réalisation d'un événement. Le joueur qui lance un dé parfaitement homogène et symétrique évalue à priori à une chance sur six la sortie d'un numéro, quel qu'il soit. Le joueur qui lance la pièce sur la table de surface S peut évaluer la probabilité que son centre tombe à l'intérieur d'une petite surface s au rapport $\frac{s}{S}$.

Nous verrons que, dans d'autres cas qui sont du domaine de la statistique, il est impossible d'évaluer raisonnablement à l'avance les chances de succès d'un résultat donné.

Voici un exemple : un enfant va naître, peut-on prévoir son sexe? La probabilité sera le résultat de l'observation. Sur 1 000 naissances, 515 sont des garçons et 485 des filles, ainsi que le constatent de

nombreuses observations. Nous évaluerons donc la probabilité pour que l'enfant soit un garçon à 0,515 et la probabilité qu'il soit une fille à 0,485.

Dans les deux cas, que la probabilité d'un événement résulte d'un raisonnement ou qu'elle soit le résultat d'une statistique, elle satisfait à une théorie axiomatique que nous allons exposer maintenant.
Soit E un ensemble fini d'événements élémentaires relatifs à une épreuve ε.

Axiome I la probabilité d'un événement A est un nombre non négatif : $\Pr(A) \geqslant 0$.

Axiome II la probabilité de l'événement certain E est égale à 1 : $\Pr(E) = 1$.

Axiome III la probabilité de A ou de B (A et B étant deux événements incompatibles) est égale à la somme des probabilités de ces événements :

$$\Pr(A \cup B) = \Pr(A) + \Pr(B).$$

Ce dernier axiome est appelé l'axiome d'additivité.

> Une urne contient 10 boules blanches, 15 boules noires et 20 boules rouges. On tire une boule de l'urne, et l'on cherche la probabilité pour qu'elle soit blanche *ou* noire.
>
> $\Pr(A) = \frac{10}{45}$: probabilité de tirer une boule blanche.
>
> $\Pr(B) = \frac{15}{45}$: probabilité de tirer une boule noire.
>
> $$\Pr(A \cup B) = \Pr(A) + \Pr(B) = \frac{10}{45} + \frac{15}{45} = \frac{25}{45} = \frac{5}{9}.$$

Ces trois premiers axiomes concernent les probabilités discrètes.

Les probabilités continues sont des nombres réels satisfaisant aux trois premiers axiomes et à l'axiome IV.

Axiome IV si $A_1, A_2, \ldots, A_n$ forment une suite dénombrable d'événements de P*, tels que tous ces événements soient incompatibles deux à deux, alors :

$$\Pr(A_1 \cup A_2 \cup \ldots \cup A_n) = \Pr(A_1) + \Pr(A_2) + \ldots + \Pr(A_n).$$

C'est l'axiome d'additivité complète.

De ces axiomes, il est facile de tirer quelques conséquences immédiates et très utiles dans la pratique :

— Si Pr (A) est la probabilité d'un événement A, la probabilité de l'événement contraire $\overline{A}$ est 1 — Pr (A). En effet, quel que soit A :

$$A \cup \overline{A} = E \quad \text{et} \quad A \cap \overline{A} = \varnothing.$$

D'où : $\Pr(A \cup \overline{A}) = \Pr(A) + \Pr(\overline{A}) = \Pr(E) = 1$ (axiome III).

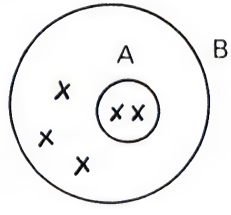
— La probabilité de l'événement impossible est nulle, puisque l'ensemble vide est le complément de E dans E.

— Pour tout événement A, Pr (A) est inférieur ou égal à 1, puisque :

$$\Pr(A) + \Pr(\overline{A}) = 1 \text{ (axiome II)}$$

et qu'une probabilité ne peut être négative.

— Si A implique B, $A \leqslant B$ et $\Pr(A) \leqslant \Pr(B)$. En effet, on voit sur le schéma ci-contre que :



$$A \cup (B \cap \overline{A}) = B$$

et : $A \cap (B \cap \overline{A}) = \varnothing.$

Par conséquent, d'après l'axiome III, les événements étant incompatibles :

$$\Pr(A \cup (B \cap \overline{A})) = \Pr(A) + \Pr(B \cap \overline{A})$$

$$\Pr(B) = \Pr(A) + \Pr(B \cap \overline{A}).$$

Et, puisque les propriétés sont toujours positives :

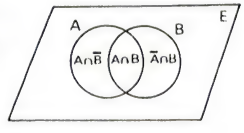
$$\Pr(A) \leqslant \Pr(B).$$

— L'axiome III peut se généraliser de la façon suivante : si $A_1, A_2, \ldots, A_n$ sont des événements incompatibles deux à deux, la probabilité de $A_1 \cup A_2 \cup \ldots \cup A_n$ est égale à la somme des probabilités :

$$\Pr(A_1) + \Pr(A_2) + \ldots + \Pr(A_n).$$

La démonstration de ce théorème se fait par la méthode de récurrence.

L'axiome III et sa généralisation ne s'appliquent qu'à des événements incompatibles. Il est utile d'envisager le cas de l'union de deux événements quelconques. Soient donc deux événements A et B qui ne sont pas incompatibles. Remarquons d'abord, en vérifiant les égalités sur la figure :



$$A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (\overline{A} \cap B).$$

Les trois événements du deuxième membre sont incompatibles deux à deux.

$$A = (A \cap \overline{B}) \cup (A \cap B).$$

Les deux événements du deuxième membre sont incompatibles. On peut donc écrire :

(1) $\Pr(A \cup B) = \Pr(A \cap \overline{B}) + \Pr(A \cap B) + \Pr(\overline{A} \cap B)$.

(2) $\Pr(A) = \Pr(A \cap \overline{B}) + \Pr(A \cap B)$.

(3) $\Pr(B) = \Pr(A \cap B) + \Pr(\overline{A} \cap B)$.

En additionnant membres à membres les égalités (2) et (3) :

$$\Pr(A) + \Pr(B) = \Pr(A \cap \overline{B}) + 2\Pr(A \cap B) + \Pr(\overline{A} \cap B).$$

Et en rapprochant cette dernière égalité de l'égalité (3) :

$$\Pr(A) + \Pr(B) = \Pr(A \cup B) + \Pr(A \cap B).$$

Cette relation se présente le plus souvent sous la forme :

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

Si A et B sont incompatibles, $\Pr(A \cap B) = 0$, et l'on retrouve l'axiome III.

Cette relation très importante se généralise au cas de plusieurs événements. Dans le cas de trois événements, elle s'écrira :

$$\Pr(A \cup B \cup C) = \Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) + \Pr(A \cap B \cap C).$$

Supposons un ensemble de 100 jetons de formes et de couleurs différentes : 40 % sont ronds, 60 % sont rouges, 25 % sont à la fois ronds et rouges. On cherche la probabilité de tirer un jeton qui soit rond *ou* rouge.
$\Pr(A) = 0{,}40$: probabilité de tirer un jeton rond.
$\Pr(B) = 0{,}60$: probabilité de tirer un jeton rouge.

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$
$$= 0{,}40 + 0{,}60 - 0{,}25 = 0{,}75.$$

Si aucun jeton rond n'était rouge, $\Pr(A \cap B) = 0$ et :

$$\Pr(A \cup B) = 0{,}40 + 0{,}60 = 1.$$

Les axiomes et les conséquences qui s'en déduisent permettent de résoudre déjà certains problèmes de probabilité, dans le cas d'un nombre fini d'événements, à condition de connaître les probabilités de chaque événement élémentaire relatif à une épreuve déterminée. L'ensemble des probabilités des événements élémentaires constitue une *base de probabilités*. Cette base permet de calculer les probabilités de n'importe quel événement relatif à cette épreuve.
La seule condition à respecter dans chaque cas est : la somme de probabilités de chaque événement élémentaire doit être égale à 1.

Les probabilités des événements élémentaires sont toutes égales.

S'il y a n événements, la probabilité de réalisation de l'un d'entre eux est $\frac{1}{n}$, puisque la somme des probabilités est égale à 1.

Soit une épreuve ε telle que :

$$E = \{ A_1, A_2, A_3, \ldots, A_n \},$$

$A_1, A_2, \ldots, A_n$ étant les événements élémentaires. Si ces derniers sont équiprobables :

$$\Pr(A_i) = \frac{1}{n}$$

et $$\Pr(A_1 \cup A_2 \cup \ldots \cup A_m) = \frac{m}{n}.$$

Or, comme nous l'avons déjà dit, tous les événements relatifs à une épreuve peuvent se réduire à l'union de plusieurs événements élémentaires ; il s'ensuit que, dans le cas de l'équiprobabilité des événements élémentaires, la probabilité d'un événement quelconque pourra s'exprimer par le rapport $\frac{m}{n}$: rapport du nombre de cas favorables à la réalisation de l'événement, à celui des cas possibles. Ces deux nombres de cas se calculeront soit par dénombrement complet, soit à l'aide de l'analyse combinatoire.

Soit, par exemple, à trouver la probabilité, pour un joueur de bridge, d'avoir les 4 as dans son jeu. Le nombre de jeux différents *pour un joueur* est égal à toutes les façons de combiner 13 à 13 les 52 cartes du jeu, soit :

$$C_{52}^{13} = \frac{52!}{13!\ 39!}\blacklozenge.$$

◆ 52! se lit « factorielle 52 » et désigne le produit des 52 premiers nombres. Voir aussi Combinaison et Factorielle.

Si un joueur possède les 4 as, les 9 autres cartes de son jeu doivent être choisies parmi les 48 cartes qui ne sont pas des as, et cela de toutes les façons possibles, soit de $C_{48}^{9} = \frac{48!}{9!\ 39!}$ façons. La probabilité d'avoir les 4 as est égale au rapport du nombre de cas favorables au nombre de cas possibles, soit :

$$\frac{\frac{48!}{9!\ 39!}}{\frac{52!}{13!\ 39!}} = \frac{48!\ 13!}{9!\ 52!} = \frac{10 \times 11 \times 12 \times 13}{49 \times 50 \times 51 \times 52} = 0{,}013.$$

Probabilités conditionnelles

Il reste à définir ce que l'on entend par probabilités conditionnelles. Il s'agit de calculer la probabilité d'un événement B, sachant que l'événement A s'est produit. Par exemple : quelle est la probabilité de tirer un as d'un jeu de 52 cartes, sachant que l'on vient d'en tirer un ?
A est l'événement : tirer un as d'un jeu de 52 cartes.
B est l'événement : tirer un as d'un jeu de 52 cartes.
B/A (lire : B si A) est l'événement : tirer un as d'un jeu de 52 cartes, sachant que l'on vient d'en tirer un.
A ∩ B est l'événement : tirer 2 as d'un jeu de 52 cartes.

$$\Pr(A) = \frac{4}{52}$$

$$\Pr(B/A) = \frac{3}{51}$$

$$\Pr(A \cap B) = \frac{C_4^2}{C_{52}^2} = \frac{\frac{4!}{2!\,2!}}{\frac{52!}{2!\,50!}} = \frac{4!\,2!\,50!}{2!\,2!\,52!} = \frac{3 \times 4}{51 \times 52}$$

$$= \frac{3}{51} \times \frac{4}{52}\blacklozenge.$$

◆ C_4^2 = toutes les façons de tirer 2 rois parmi les 4 rois du jeu (nombre d'événements favorables).
C_{52}^2 = toutes les façons de tirer 2 cartes d'un jeu de 52 (nombre d'événements possibles).

On constate sur cet exemple :

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B/A).$$

Nous admettrons que cette relation est générale.
Dans le cas particulier où la probabilité de l'événement B n'est pas influencée par la réalisation de l'événement A, on dit que *les deux événements sont indépendants*. On a alors :

$$\Pr(B/A) = \Pr(B),$$

et $$\Pr(A \cap B) = \Pr(A) \times \Pr(B).$$

Par exemple, si, après avoir sorti un as d'un jeu de 52 cartes (événement A), on remet cet as dans le jeu et l'on désire sortir encore un as (événement B) :

$$\Pr(A) = \frac{4}{52}$$

$$\Pr(B) = \frac{4}{52}$$

$$\Pr(B/A) = \Pr(B) = \frac{4}{52}.$$

La probabilité pour que, dans ces conditions de jeu, on sorte 2 as est :

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B) = \frac{4}{52} \times \frac{4}{52}.$$

Incompatibilité, indépendance

Il importe de ne pas confondre événements incompatibles et événements indépendants. Incompatibilité signifie que les deux résultats ne peuvent être obtenus simultanément et indépendance signifie que la probabilité de l'un n'est pas modifiée par la réalisation de l'autre.
On jette deux dés et l'on considère les deux événements : sortir un 2 avec le premier dé et faire un total de 10 après le jet du deuxième dé. Voilà deux événements incompatibles.
On jette deux dés et l'on considère les deux événements : sortir un 2 avec le premier dé et faire un total supérieur à 2 après le jet du deuxième dé. Voilà deux événements indépendants.

VARIABLE ALÉATOIRE

Il arrive très souvent que le résultat d'un événement puisse s'exprimer par un nombre. Par exemple : le numéro sorti en jetant un dé, le nombre de « piles » obtenues en jetant dix fois une pièce de monnaie, le nombre de coups de téléphone reçus par une entreprise en une heure. On conçoit facilement que l'on puisse remplacer l'événement par ce nombre. Mais il peut arriver aussi que les événements élémentaires d'une épreuve ne s'expriment pas par des nombres. Il est alors toujours possible d'établir conventionnellement une correspondance entre ces événements et un ensemble des nombres. Par exemple : l'ensemble E des événements élémentaires relatifs à l'épreuve « jeter une pièce de monnaie » s'écrit :

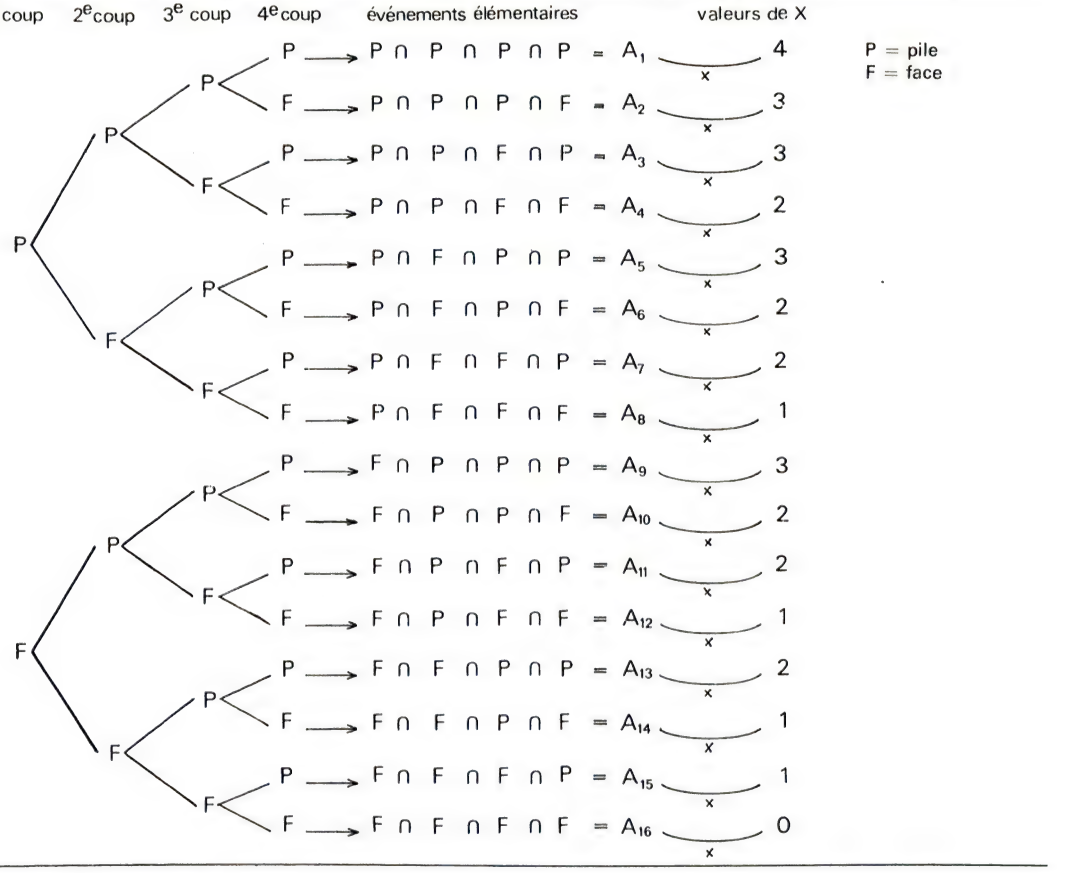
$$E = \{ P, F \}.$$

On fera correspondre le nombre 1 au résultat « pile » et le nombre 0 au résultat « face ». L'ensemble E peut se remplacer par le nouvel ensemble $\{ 1, 0 \}$. De même, s'il s'agit du sexe d'un nouveau-né, on peut remplacer « garçon » par 1 et « fille » par 0.

On appellera variable aléatoire une application de l'ensemble des événements élémentaires sur $\mathbb{R}$: *ensemble des nombres réels.* On désignera les variables aléatoires par une majuscule romaine et les valeurs prises par les applications dans $\mathbb{R}$ par des minuscules.
Soit X cette application ; l'ensemble des images des événements élémentaires par X constitue l'ensemble des valeurs de la variable aléatoire.

Par exemple : jetons quatre fois une pièce de monnaie. Pour définir l'ensemble E, nous ferons un arbre, et nous appellerons X le nombre de piles sorties.

L'ensemble des valeurs de la variable aléatoire est : $\{ 0, 1, 2, 3, 4 \}$.



LOIS DE PROBABILITÉ

La notion de variable aléatoire permet de substituer un sous-ensemble de $\mathbb{R}$ à l'ensemble des événements élémentaires. C'est ainsi que, dans l'exemple précédent, nous remplacerons l'ensemble E par l'ensemble des valeurs de X, soit $\{ 0, 1, 2, 3, 4 \}$. A chacune de ces valeurs va s'associer une probabilité. Ce sera la probabilité de l'image réciproque par X.

Valeurs de la variable : x_i l'ensemble des valeurs étant $\{0, 1, 2, 3, 4\}$	Images réciproques associées à la variable X (voir tableau précédent).	Probabilités : p_i
$x_1 = 0 \longleftrightarrow$	A_{16}	$\Pr(A_{16}) = \Pr(X = x_1) = \frac{1}{16}$.
$x_2 = 1 \longleftrightarrow$	$A_8 \cup A_{12} \cup A_{14} \cup A_{15}$	$\Pr(A_8) + \Pr(A_{12}) + \Pr(A_{14}) + \Pr(A_{15}) = \Pr(X = x_2) = \frac{4}{16}$.
$x_3 = 2 \longleftrightarrow$	$A_4 \cup A_6 \cup A_{10} \cup A_{11} \cup A_{13}$	$\Pr(A_4) + \Pr(A_6) + \Pr(A_{10}) + \Pr(A_{11}) + \Pr(A_{13}) = \Pr(X = x_3) = \frac{6}{16}$.
$x_4 = 3 \longleftrightarrow$	$A_2 \cup A_3 \cup A_5 \cup A_9$	$\Pr(A_2) + \Pr(A_3) + \Pr(A_5) + \Pr(A_9) = \Pr(X = x_4) = \frac{4}{16}$.
$x_5 = 4 \longleftrightarrow$	A_1	$\Pr(A_1) = \Pr(X = x_5) = \frac{1}{16}$.

On remarque sur cet exemple que :

$$\Pr(X = x_i) = p_i \leqslant 1$$

$$\text{et : } \sum_{i=1}^{n} \Pr(X = x_i) = \sum_{i=1}^{n} p_i$$

$$\text{soit } \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = 1.$$

L'ensemble des couples (x_i, p_i) *s'appelle loi de probabilité de la variable* X.

Nous distinguerons :

a) Les variables aléatoires discrètes finies, dont le nombre des valeurs est fini. C'est le cas de l'exemple précédent. La loi de probabilité sera parfaitement définie par l'ensemble des couples (x_i, p_i). La somme des probabilités relatives à chaque valeur de la variable est égale à 1.

b) Les variables aléatoires discrètes infinies dont le nombre des valeurs est dénombrable mais infini. Dans ce cas-là, on devra établir une relation mathématique entre les deux éléments du couple (x_i, p_i). Reprenons l'exemple précédent et jetons la pièce n fois. La probabilité pour que pile sorte k fois est égale à :

$$p_k = C_n^k \left(\frac{1}{2}\right)^n \blacklozenge.$$

◆ Ce résultat est établi lors de l'étude de la loi binômiale (voir l'article « les Statistiques », page 486).

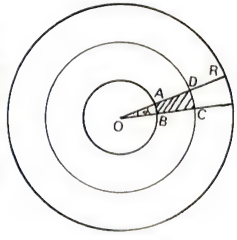
La loi de probabilité est l'ensemble des couples :

$$\left(x_{k+1},\ C_n^k \left(\frac{1}{2}\right)^n\right)$$

pour k variant de 0 à n. On a encore :

$$\sum_{i=1}^{n} p_i = 1.$$

c) Les variables aléatoires continues dont l'ensemble des valeurs A est continu. Par exemple, lorsqu'on jette des flèches sur une cible, le point d'impact est une variable aléatoire continue. Soit S la surface de la cible $= \pi R^2$. Soit s la surface d'une portion de couronne ABCD comprise entre les deux cercles de rayons r_1 et r_2. L'angle $\widehat{AOB} = \alpha$; $s = \frac{\alpha}{2}(r_1^2 - r_2^2)$. La probabilité pour que la flèche tombe dans la surface ABCD est égale à :



$$\frac{s}{S} = \frac{\alpha\,(r_1^2 - r_2^2)}{2\,\pi\,R^2}.$$

Si cette surface diminue jusqu'à se réduire à un point, r_2 tend vers r_1 et s tend vers 0, et, par conséquent, la probabilité pour que la flèche tombe en un point donné est nulle.

Fonction de répartition

Dans ces conditions, il est impossible de définir la loi de probabilité pour l'ensemble des couples (x_i, p_i). On introduit alors la notion de fonction de répartition de la variable aléatoire X :

$$F(x) = Pr(X < x).$$

C'est la probabilité pour que les valeurs prises par la variable aléatoire soient inférieures à une valeur donnée.

Densité de probabilité ◆

◆ Lire d'abord Dérivée intégrale et primitive.

Supposons une variable aléatoire continue dont les valeurs varient de $-\infty$ à $+\infty$. Si, pour toutes les valeurs de X, la fonction F (x) possède une dérivée $f(x)$ telle que :

$$F(x_0) = \int_{-\infty}^{x_0} f(x)\,dx,$$

la fonction $f(x)$ est appelée densité de probabilité de la variable aléatoire X. C'est la probabilité pour que X prenne une valeur comprise entre x et $x + dx$.

La fonction de répartition est très utilisée en statistique et il est important de noter un certain nombre de remarques essentielles.
Supposons une variable aléatoire continue X, variant dans l'intervalle a, b : $F(a) = 0$, puisque c'est la probabilité pour X d'être inférieure à a ; $F(b) = 1$, puisque c'est la probabilité pour X d'être inférieure à b et, par conséquent, de prendre toutes les valeurs possibles. En effet, la fonction étant continue :

$$\lim_{\varepsilon \to 0} F(b - \varepsilon) = F(b).$$

Cela est une conséquence de l'axiome d'additivité complète.
La fonction $F(x)$ ne peut jamais décroître ; cela résulte de ce que les probabilités sont des nombres toujours positifs ou nuls.

Fonction cumulative

La fonction de répartition peut aussi servir à décrire les lois de probabilités des variables discrètes. On l'appelle souvent fonction cumulative ◆.

◆ Dans le cas des variables aléatoires discrètes, il arrive quelquefois que l'on définisse la fonction de répartition comme la probabilité pour X de prendre une valeur inférieure ou égale à x_n.

La probabilité pour que $X < x_k$ est égale à $\sum_{i=1}^{k-1} p_i$. Lorsqu'on jette la pièce de monnaie quatre fois, si X est le nombre de fois où sort pile, la fonction de répartition prend les valeurs suivantes ◆ :

X	p_i	$F(x)$
$x_1 = 0$	$\frac{1}{16}$	$F(x_1) = 0$
$x_2 = 1$	$\frac{4}{16}$	$F(x_2) = \frac{1}{16}$
$x_3 = 2$	$\frac{6}{16}$	$F(x_3) = \frac{5}{16}$
$x_4 = 3$	$\frac{4}{16}$	$F(x_4) = \frac{11}{16}$
$x_5 = 4$	$\frac{1}{16}$	$F(x_5) = \frac{15}{16}$

◆ Voir tableau page 426.

Dans le cas des variables aléatoires discrètes, il arrive quelquefois que l'on définisse la fonction de répartition comme la probabilité pour X de prendre une valeur inférieure ou égale à x_i.

Espérance mathématique, valeur moyenne

On appelle espérance mathématique ou valeur moyenne d'une variable aléatoire la quantité E (X) définie par :

(1) $$E(X) = \sum_{i=1}^{n} x_i p_i \quad \text{pour une variable discrète}$$

(2) $$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx \quad \text{pour une variable continue.}$$

Cela suppose naturellement que la suite (1) comme l'intégrale (2) soient convergentes.
La valeur moyenne de la variable aléatoire X : nombre de fois où pile est sortie lorsqu'on jette quatre fois une pièce, est :

$$E(X) = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16} = 2.$$

On peut citer un cas classique où l'espérance mathématique ne peut se calculer ; c'est le fameux « paradoxe de Saint-Pétersbourg ». A et B jouent à pile ou face. B s'engage à verser 1 franc à A si face sort au premier coup ; 2 francs, si face sort au deuxième coup pour la première fois ; 2^2 francs, si face sort au troisième coup pour la première fois, etc. Le jeu s'arrête lorsque face est sortie. Quelle doit être la mise de A pour que le jeu soit équitable? Notons en passant qu'un jeu est équitable si la mise d'un joueur est égale à son espérance mathématique. On entend par espérance mathématique d'un joueur le produit de son gain par la probabilité de le gagner. Il s'agit donc bien de calculer la valeur moyenne d'une variable aléatoire X qui est le gain du joueur.

valeurs de X	probabilités	$x_i\ p_i$
1	$\frac{1}{2}$	$\frac{1}{2}$
2	$\frac{1}{2} \times \frac{1}{2}$	$\frac{1}{2}$
2^2	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$	$\frac{1}{2}$
2^3	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$	$\frac{1}{2}$
etc		

$$E(X) = \sum_{i=1}^{\infty} x_i p_i = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$$

E (X) est infinie. La mise de A devrait donc être infinie. En réalité, ce jeu n'est un paradoxe que parce que l'on suppose la fortune de B également infinie. Mais, si l'on sait que, pour lancer la pièce 25 fois, il faudrait que B dispose de plus de 33 millions, on se rend compte que ce jeu n'est possible qu'en fixant à l'avance le nombre de coups en fonction de la fortune de B. Dans ce cas, l'espérance mathématique de A peut se calculer.

Ecarts

La notion d'écart, très utilisée en statistique, consiste à mesurer les valeurs de la variable aléatoire par rapport à une nouvelle origine.

Ainsi, en 1972, au lieu de dire qu'un monsieur a 50 ans, on peut dire qu'il est de la classe 42, ce qui signifie qu'en prenant pour origine l'âge de 20 ans ce monsieur présente un écart de 30 ans par rapport à cette nouvelle origine.
En particulier, nous pouvons mesurer toutes les valeurs d'une variable aléatoire X en prenant comme origine sa valeur moyenne. On obtient un nouvel ensemble de valeurs appelées écarts de X et égaux à $y_i = x_i - \mathrm{E}(X)$. Cette nouvelle variable aléatoire Y présente une caractéristique très importante : sa valeur moyenne est nulle. En effet :

$$\mathrm{E}(Y) = \sum_{i=1}^{n} y_i p_i = \sum_{i=1}^{n} \left(x_i - \mathrm{E}(X)\right) p_i$$

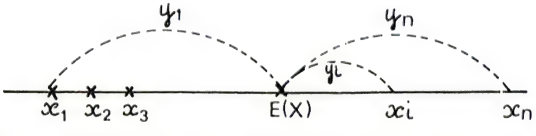
$$= \sum_{i=1}^{n} x_i p_i - \Sigma(X) \sum_{i=1}^{n} p_i = \mathrm{E}(X) - \mathrm{E}(X), \text{ puisque } \sum_{i=1}^{n} p_i = 1$$

$$\mathrm{E}(Y) = 0.$$

La démonstration serait la même dans le cas d'une variable continue.

Variance, écart type

La nouvelle variable Y a pour intérêt primordial de nous renseigner sur la dispersion des valeurs de la variable aléatoire X autour de sa moyenne. En effet, plus les grandes valeurs de Y sont nombreuses, plus les valeurs de X « s'étalent » de part et d'autre de sa valeur moyenne :



Il serait donc intéressant de représenter les valeurs de Y par un paramètre dont la grandeur renseignerait sur la dispersion des valeurs de X ; E (Y) étant nulle, on calcule E (Y^2), c'est-à-dire la valeur moyenne des carrés des valeurs de Y.
E (Y^2) se nomme la variance, $\sqrt{\mathrm{E}(Y^2)}$ se nomme l'écart quadratique moyen ou écart type ; il est souvent désigné par σ.

x_i	$y_i = x_i - 2$	y_i^2	$y_i^2 p_i^2$
0	−2	4	$\frac{4}{16}$
1	−1	1	$\frac{4}{16}$
2	0	0	0
3	1	1	$\frac{4}{16}$
4	2	4	$\frac{4}{16}$
			1

$$y_i = x_i - \mathrm{E}(X).$$

Dans notre exemple, E(X) = 2 (voir page précédente).
Calculons l'écart type de la variable X : nombre de fois où sort pile en jetant quatre fois une pièce de monnaie.

$$\Sigma\, y_i^2 p_i = 1.$$

$$\sigma = \sqrt{1} = 1.$$

L'écart type σ est une caractéristique essentielle de la répartition des valeurs d'une variable aléatoire.

Si $\sigma = 0$, toutes les valeurs de Y sont nulles et toutes les valeurs de X sont égales à sa valeur moyenne. X n'est plus une variable aléatoire.

Si σ est petit, c'est qu'il y a peu de grandes valeurs de Y ou, s'il y en a beaucoup, elles sont peu probables.

Moments

La valeur moyenne, l'écart type ne sont que deux « moments » particuliers. On appelle moment d'ordre k la valeur moyenne des valeurs de X élevées à la puissance k :

$$\mathrm{E}(\mathrm{X}^k) = \int_{-\infty}^{+\infty} x^k f(x)\,\mathrm{d}x$$

ou : $$\mathrm{E}(\mathrm{X}^k) = \sum_{i=1}^{n} x_i^k p_i.$$

On appelle moment centré d'ordre k la valeur moyenne des valeurs de Y élevées à la puissance k :

$$\mathrm{E}(\mathrm{Y}^k) = \int_{-\infty}^{+\infty} (x - \mathrm{E}(\mathrm{X}))^k f(x)\,\mathrm{d}x$$

ou : $$\mathrm{E}(\mathrm{Y}^k) = \sum_{i=1}^{n} (x_i - \mathrm{E}(\mathrm{X}))^k p_i.$$

Le moment centré d'ordre 1 est nul. Le moment centré d'ordre 2 est l'écart type de X. Les moments d'ordre supérieur à 2 donnent des indications sur certaines caractéristiques de forme de la fonction de répartition, telles que : asymétrie, aplatissement.

Inégalité de Bienaymé-Tchebycheff

L'écart type σ donne une idée de la dispersion des valeurs de la variable aléatoire autour de sa moyenne. L'inégalité de Bienaymé-Tchebycheff ◆ exprime que, pour tout λ appartenant à $\mathbb{R}^+$ (ensemble des nombres réels positifs), la probabilité pour que la valeur absolue de Y soit supérieure ou égale à $\lambda\sigma$ est inférieure ou égale à $\frac{1}{\lambda^2}$. Nous allons établir cette relation dans le cas d'une variable continue. Elle est également valable pour une variable discrète.
Un écart en valeur absolue supérieur ou égal à $\lambda\sigma$ peut se réaliser

◆ Bienaymé : statisticien français (1796-1878) ; Tchebycheff : mathématicien russe (1821-1894).

de deux manières incompatibles : ou bien l'écart est positif et supérieur à $\lambda\sigma$, ou bien l'écart est négatif et inférieur à $\lambda\sigma$.

D'où : $\Pr(|Y| \geqslant \lambda\sigma) = \Pr(Y \leqslant -\lambda\sigma) + \Pr(Y \geqslant \lambda\sigma)$.

Soit $f(y)$ la densité de probabilité de la variable Y :

$$\sigma^2 = \int_{-\infty}^{+\infty} y^2 f(y)\,dy = \int_{-\infty}^{-\lambda\sigma} y^2 f(y)\,dy + \int_{-\lambda\sigma}^{+\lambda\sigma} y^2 f(y)\,dy + \int_{+\lambda\sigma}^{+\infty} y^2 f(y)\,dy.$$

Il s'ensuit que :

$$(1) \qquad \sigma^2 \geqslant \int_{-\infty}^{-\lambda\sigma} y^2 f(y)\,dy + \int_{+\lambda\sigma}^{+\infty} y^2 f(y)\,dy,$$

puisque f est une fonction non négative.

Par ailleurs, comme $|Y| \geqslant \lambda\sigma$, $Y^2 \geqslant \lambda^2\sigma^2$; on peut remplacer, dans l'inégalité (1), y^2 par $\lambda^2\sigma^2$, ce qui donne :

$$\sigma^2 \geqslant \lambda^2\sigma^2 \left(\int_{-\infty}^{-\lambda\sigma} f(y)\,dy + \int_{+\lambda\sigma}^{+\infty} f(y)\,dy\right)$$

ou encore :

$$\sigma^2 \geqslant \lambda^2\sigma^2\,[\Pr(Y < -\lambda\sigma) + \Pr(Y > \lambda\sigma)],$$

soit :

$$\sigma^2 \geqslant \lambda^2\sigma^2 \times \Pr(|Y| > \lambda\sigma)$$

et, par conséquent :

$$\Pr(|Y| > \lambda\sigma) \leqslant \frac{1}{\lambda^2}.$$

En fait, la probabilité d'un écart supérieur en valeur absolue à $\lambda\sigma$ est souvent bien inférieure à $\frac{1}{\lambda^2}$. Le lemme de Bienaymé nous servira plus tard à démontrer le théorème de Bernoulli, qui sert de beaucoup plus près la véritable valeur de cette probabilité. Pour le moment, l'inégalité de Bienaymé-Tchebycheff a l'intérêt de montrer comment l'écart type σ intervient dans la mesure de la dispersion de la série.

Couples aléatoires

La notion de variable aléatoire peut se généraliser. En particulier, il est possible de considérer le cas où cette variable est un couple

(X, Y), l'ensemble des valeurs de la variable étant une partie de $\mathbb{R}^2$ ◆. Pratiquement, le couple aléatoire correspond aux événements d'une épreuve comprenant deux groupes de résultats. Par exemple, un joueur de cartes peut considérer son jeu sous l'angle de la « force des cartes » et sous l'angle de leur « couleur ». L'événement : obtenir un as de pique, est le couple (as, P). La probabilité d'obtenir un as de pique est :

◆ Ensemble des couples du produit cartésien $\mathbb{R} \times \mathbb{R}$.

$$\Pr(\text{as, P}) = \Pr[(X = \text{as}) \cap (Y = \text{P})].$$

Nous n'entreprendrons pas l'étude théorique des couples aléatoires.

Cette partie, un peu abstraite, débouche sur des applications pratiques très importantes. Nous la retrouvons sous une forme concrète dans l'étude de la corrélation ◆.

◆ Voir l'article « les Statistiques ».

Somme de plusieurs variables aléatoires

On est souvent amené à considérer une variable aléatoire comme la somme de plusieurs autres. Par exemple : on jette un dé blanc à quatre faces et un dé rouge à quatre faces. Soient X le point obtenu avec le dé blanc et Y le point obtenu avec le dé rouge.

$Z = X + Y$ représentera le total des points ; Z est une variable aléatoire qui est elle-même la somme de deux variables aléatoires.

Valeur moyenne (ou espérance mathématique) d'une somme de variables aléatoires indépendantes

Proposons-nous de calculer la valeur moyenne de Z en fonction des valeurs moyennes de X et de Y, en remarquant que X et Y sont deux variables aléatoires indépendantes.

Loi de probabilité de X : lorsque la valeur de X est 1, $\Pr(X = x_i) = \frac{1}{4}$; lorsque $X = 2$, $\Pr(X = x_i) = \frac{1}{4}$; de même pour $X = 3$ et $X = 4$.

$$E(X) = 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} + 4 \times \frac{1}{4} = \frac{10}{4} = 2{,}5.$$

La loi de probabilité de Y est la même que celle de X et son espérance mathématique aussi :

$$E(Y) = \frac{10}{4} = 2{,}5.$$

La loi de probabilité de Z s'obtient de la façon suivante◆ :

ensemble des événements possibles	probabilité de chaque événement $\Pr(X = x_i) \times \Pr(Y = y_i)$
1+1 = 2	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
1+2 = 3	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
1+3 = 4	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
1+4 = 5	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
2+1 = 3	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
2+2 = 4	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
2+3 = 5	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
2+4 = 6	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
3+1 = 4	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
3+2 = 5	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
3+3 = 6	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
3+4 = 7	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
4+1 = 5	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
4+2 = 6	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
4+3 = 7	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$
4+4 = 8	$\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$

loi de probabilité de Z

valeurs de Z	$\Pr(Z = z_i)$
2	$\frac{1}{16}$
3	$\frac{2}{16}$
4	$\frac{3}{16}$
5	$\frac{4}{16}$
6	$\frac{3}{16}$
7	$\frac{2}{16}$
8	$\frac{1}{16}$

◆ Dans la partie gauche du tableau, on envisage tous les cas possibles en jetant les deux dés. A chaque cas correspond une valeur de Z. On évalue la probabilité de chacun de ces événements en faisant le produit des probabilités des deux événements composants (voir « Probabilité conditionnelle, cas de deux événements indépendants », page 423). De cette partie gauche du tableau on déduit la loi de probabilité de Z (partie droite du tableau). Exemple : Z peut prendre la valeur 4 de trois manières : événement A : 1 + 3, événement B : 2 + 2, événement C : 3 + 1.

$\Pr(Z = 3)$

$= \Pr(A \cup B \cup C)$

$= \Pr(A) + \Pr(B) + \Pr(C)$

$= \frac{3}{16}$ (axiome d'additivité).

$$E(Z) = 2 \times \frac{1}{16} + 3 \times \frac{2}{16} + 4 \times \frac{3}{16} + 5 \times \frac{4}{16} + 6 \times \frac{3}{16} + 7 \times \frac{2}{16} + 8 \times \frac{1}{16} = 5.$$

Nous remarquons sur cet exemple que : $E(Z) = E(X) + E(Y)$.
Nous admettrons ce résultat dans le cas général, lequel s'étendrait facilement à la somme de plusieurs variables : *la valeur moyenne d'une somme de plusieurs variables aléatoires indépendantes est égale à la somme des valeurs moyennes de chaque variable aléatoire.*

Variance d'une somme de variables aléatoires indépendantes

Proposons-nous de calculer la variance de Z : σ_Z^2, en fonction des variances de X : σ_X^2 et de Y : σ_Y^2 :

$$\sigma_X^2 = \Sigma(x_i - E(X))^2 \Pr(X = x_i)$$

$$= (1 - 2{,}5)^2 \times \frac{1}{4} + (2 - 2{,}5)^2 \times \frac{1}{4} + (3 - 2{,}5)^2 \times \frac{1}{4}$$

$$+ (4 - 2{,}5)^2 \times \frac{1}{4} = \frac{5}{4}.$$

$$\sigma_Y^2 = \frac{5}{4}.$$

$$\sigma_Z^2 = (2-5)^2 \times \frac{1}{16} + (3-5)^2 \times \frac{2}{16} + (4-5)^2 \times \frac{3}{16} + (5-5)^2 \times \frac{4}{16} + (6-5)^2 \times \frac{3}{16} + (7-5)^2 \times \frac{2}{16} + (8-5)^2 \times \frac{1}{16} = \frac{10}{4}.$$

On constate donc que :

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2.$$

Nous admettrons ce résultat dans le cas général, lequel s'étendrait facilement à la somme de plusieurs variables : *la variance de la somme de plusieurs variables aléatoires indépendantes est égale à la somme des variances de chaque variable aléatoire.*

Variables aléatoires non indépendantes

La valeur moyenne d'une somme de variables aléatoires non indépendantes est encore égale à la somme des valeurs moyennes de chaque variable aléatoire.
Mais il n'en est pas de même de la variance. On obtiendrait, dans le cas de deux variables aléatoires dépendantes X et Y susceptibles de prendre respectivement n et m valeurs :

$$\sigma_{(X+Y)}^2 = \sigma_X^2 + \sigma_Y^2 + 2 \sum_{i=1}^{i=n} \sum_{j=1}^{j=m} (x_i - E(X))(y_j - E(Y)).$$

La double sommation $\sum_{i=1}^{i=n} \sum_{j=1}^{j=m} (x_i - E(X))(y_j - E(Y))$ se nomme covariance. Elle est nulle dans le cas de deux variables indépendantes.

LOI DES GRANDS NOMBRES

Reprenons l'exemple des dés tétraédriques et jetons n dés semblables. Soient $X_1, X_2, \ldots, X_n$ n variables aléatoires prenant *chacune* les valeurs 1, 2, 3, 4 (données par le jet de chaque dé). Ces variables obéissent toutes à la même loi de probabilité (voir page 425) dont la valeur moyenne est $E(X) = 2{,}5$ ◆. Soit S_n la variable aléatoire suivante :

◆ Voir ci-dessus, page 433.

$$S_n = X_1 + X_2 + \ldots + X_n$$

$$\frac{S_n}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Les valeurs de $\frac{S_n}{n}$ sont comprises entre $\frac{n}{n} = 1$ et $\frac{4n}{n} = 4$.

Lorsque n augmente indéfiniment, les valeurs de $\frac{S_n}{n}$ se rapprochant de E (X) = 2,5 ont une probabilité de plus en plus grande, qui tend vers 1. On dit que $\frac{S_n}{n}$ converge en probabilité vers E (X). C'est la loi des grands nombres ; elle peut s'exprimer ainsi :

$$\lim_{n \to \infty} \Pr\left(\left| \frac{S_n}{n} - E(X) \right| < \varepsilon\right) = 1$$

pour ε arbitrairement petit.

Démonstration. Pour le démontrer, nous utiliserons l'inégalité de Bienaymé-Tchebycheff ◆.

◆ Voir page 431.

$$\Pr\left(\left| \frac{S_n}{n} - E(X) \right| > \lambda\sigma\right) \leqslant \frac{1}{\lambda^2},$$

ce qui peut s'écrire :

$$\Pr\left(\left| \frac{S_n}{n} - E(X) \right| < \lambda\sigma\right) > 1 - \frac{1}{\lambda^2}.$$

Posons $\lambda\sigma = \varepsilon$, d'où $\lambda = \frac{\varepsilon}{\sigma}$.

$$(1) \qquad \Pr\left(\left| \frac{S_n}{n} - E(X) \right| < \varepsilon\right) > 1 - \frac{\sigma^2}{\varepsilon^2}.$$

σ^2 est la variance de $\frac{S}{n}$.

$$\sigma^2_{\frac{S}{n}} = E\left(\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right)^2 = E\left(\frac{1}{n} \cdot S_n - \frac{1}{n} E(S_n)\right)^2$$

$$= \frac{1}{n^2} E\left(S_n - E(S_n)\right)^2 = \frac{1}{n^2} \sigma^2_{S_n}$$

$$= \frac{1}{n^2} \times n\sigma^2_X = \frac{\sigma^2_X}{n}.$$

En reportant cette valeur dans l'inégalité (1) :

$$\Pr\left(\left| \frac{S_n}{n} - E(X) \right| < \varepsilon\right) > 1 - \frac{\sigma^2_X}{n\varepsilon^2}.$$

Si l'on fait tendre n vers ∞, on obtient à la limite :

$$\lim_{n \to \infty} \Pr\left(\left| \frac{S_n}{n} - E(X) \right| < \varepsilon \right) > 1.$$

Mais, comme une probabilité ne peut être supérieure à 1, on a ainsi prouvé le théorème.

Théorème de Bernoulli

C'est un cas particulier de la loi des grands nombres. Supposons une épreuve qui ne comporte que deux éventualités, par exemple : jeter une pièce de monnaie. La variable aléatoire X sera égale à 1 si la pièce montre pile et à 0 dans le cas contraire. Répétons cette épreuve n fois. Pile va sortir un certain nombre de fois. Nous appellerons f_n la fréquence relative de cette sortie, qui est égale au nombre de piles sur le nombre de coups. Mais f_n est aussi égale à :

$$f_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

puisque X_i, qui est le résultat du i-ième coup, est égal à 1 ou à 0 selon que pile est sortie ou non. D'après la loi des grands nombres, lorsque n augmente, f_n se rapproche de E(X). Calculons E(X) :

$$E(X) = 1 \times \frac{1}{2} + 0 \times \frac{1}{2} = \frac{1}{2}.$$

Dans le cas général d'une variable aléatoire égale à 1 avec une probabilité p et à 0 avec une probabilité $1 - p$:

$$E(X) = 1 \times p + 0\,(1 - p) = p.$$

Nous pouvons donc énoncer le résultat : si un événement A est le résultat aléatoire d'une épreuve et si l'on répète indéfiniment cette même épreuve, la fréquence relative d'arrivée de l'événement A converge en probabilité vers sa propre probabilité :

$$\lim_{n \to \infty} \Pr(|f_n - p| < \varepsilon) = 1$$

pour ε arbitrairement petit.

Ce résultat est connu depuis Jacques Bernoulli ◆.

◆ J. Bernoulli : mathématicien suisse (1654-1705).

Remarquons bien que le théorème de Bernoulli n'affirme pas que $\frac{S_n}{n}$ tend vers p quand n tend vers ∞, mais seulement que la probabilité d'avoir $\left| \frac{S_n}{n} - p \right| < \varepsilon$ tend vers 1 lorsque n tend vers ∞.

Michèle Pascal.

PRODUIT CARTÉSIEN (de deux ensembles) : voir Cartésien
et les pages 274, 364.

PRODUIT REMARQUABLE : voir Binôme

PRODUIT SCALAIRE◆ pages 95 et 96, 329.

◆ Pour le lecteur peu familiarisé, il est conseillé de lire d'abord l'article « l'Algèbre linéaire ».

Si E est un espace vectoriel sur le corps $\mathbb{R}$ des nombres réels, on appelle produit scalaire sur E toute application qui fait correspondre au couple de vecteurs $(\vec{u}, \vec{v})$ un nombre réel noté $\vec{u} \cdot \vec{v}$ (ou $< \vec{u}, \vec{v} >$ ◆) et qui vérifie les propriétés suivantes :

◆ Pour cette notation voir l'article « l'Algèbre linéaire », page 95.

— pour tout $\vec{u}$ et tout $\vec{v}$ de E :

$$\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u} \text{ (commutativité)} ;$$

— pour tout $\vec{u}$, tout $\vec{u'}$ et tout $\vec{v}$ de E :

$$(\vec{u} + \vec{u'}) \cdot \vec{v} = \vec{u} \cdot \vec{v} + \vec{u'} \cdot \vec{v} \text{ (distributivité à droite)} ;$$

— pour tout $\vec{u}$, tout $\vec{v}$ et tout $\vec{v'}$ de E :

$$\vec{u} \cdot (\vec{v} + \vec{v'}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{v'} \text{ (distributivité à gauche)} ;$$

— pour tout $\vec{u}$ et tout $\vec{v}$ de E, pour tout λ de $\mathbb{R}$:

$$\vec{u} \cdot (\lambda \vec{v}) = (\lambda \vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v}) ;$$

— pour tout $\vec{u}$ de E : $\vec{u} \cdot \vec{u} \geqslant 0$◆.

◆ Le produit scalaire de $\vec{u}$ par lui-même est appelé carré scalaire de $\vec{u}$.

Exemple

Si l'espace vectoriel considéré est l'ensemble $\mathbb{R}^2$ des couples (x, y) de nombres réels, on peut vérifier que l'application qui aux deux couples (x, y) et (x', y') fait correspondre le nombre réel $xx' + yy'$ est un produit scalaire.

Deux vecteurs $\vec{u}$ et $\vec{v}$ sont orthogonaux si et seulement si leur produit scalaire est nul. Une base de l'espace vectoriel E constituée de vecteurs orthogonaux deux à deux est appelée base orthogonale.

L'application qui à tout vecteur $\vec{u}$ de E fait correspondre le nombre réel positif $\sqrt{\vec{u} \cdot \vec{u}}$ est une norme sur E. Un vecteur $\vec{u}$ dont la norme est égale à 1 est un vecteur unitaire.
Une base de E formée de vecteurs orthogonaux deux à deux et unitaires est une base orthonormée.

Si E est de dimension finie n, si $\vec{u}$ a pour composantes $(x_1, x_2, \ldots, x_n)$ dans une base orthonormée de E et si $\vec{v}$ a pour composantes $(y_1, y_2, \ldots, y_n)$ dans la même base, l'expression du produit scalaire $\vec{u} \cdot \vec{v}$ est : $x_1y_1 + x_2y_2 + \ldots + x_ny_n$.

On peut définir le cosinus de l'angle des deux vecteurs $\vec{u}$ et $\vec{v}$ par la formule :

$$\vec{u} \cdot \vec{v} = \left\| \vec{u} \right\| \times \left\| \vec{v} \right\| \times \cos (\vec{u}, \vec{v}),$$

où $\left\| \vec{u} \right\|$ et $\left\| \vec{v} \right\|$ sont les normes des vecteurs $\vec{u}$ et $\vec{v}$ respectivement.

PROJECTIF : voir Géométrie et Coordonnées homogènes

PROJECTION

Dans le plan, étant donné deux droites non parallèles D et Δ, on appelle projection sur D parallèlement à Δ la transformation qui à tout point M du plan fait correspondre le point M′, intersection de D avec la parallèle à Δ passant par M (fig. 1). Si la droite D et la droite Δ sont perpendiculaires, on dit que la projection est orthogonale (fig. 2).

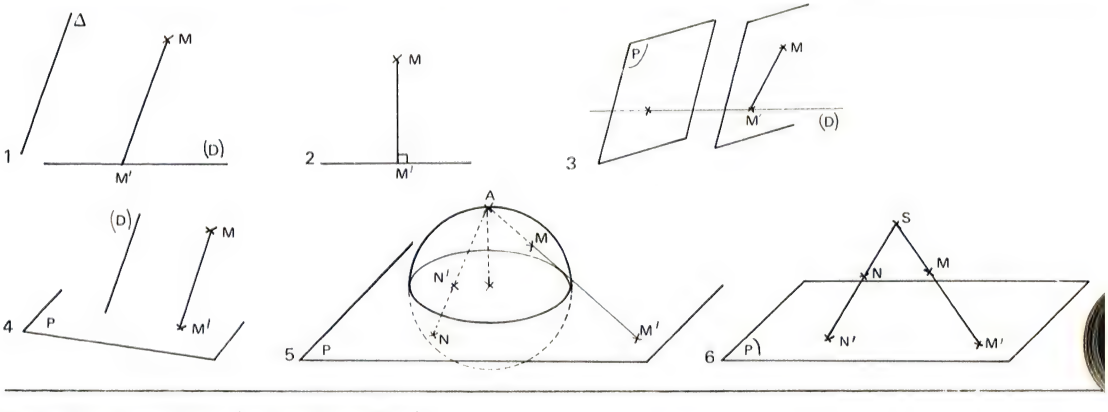
Dans l'espace, étant donné un plan P et une droite non parallèle, on peut définir deux projections :

— *projection sur* D *parallèlement à* P : transformation qui à chaque point M fait correspondre le point M′, intersection de D avec le plan parallèle à P passant par M (fig. 3) ;

— *projection sur* P *parallèlement à* D : transformation qui à chaque point M fait correspondre le point M′, intersection de P avec la droite parallèle à D passant par M (fig. 4).

Projection stéréographique

Si (S) est une sphère, P un plan passant par le centre de cette sphère et A un des deux points de la sphère où le plan tangent est parallèle à P, on appelle projection stéréographique la transformation géométrique qui à tout point M de la sphère fait correspondre le point M′, intersection de la droite AM avec le plan P (fig. 5).



Projection centrale (ou perspective)

Étant donné un point S de l'espace et un plan P ne contenant pas S, la projection centrale de sommet S est la transformation qui à tout point (autre que S) de l'espace fait correspondre le point M′, intersection de la droite SM avec le plan P (fig. 6).

PUISSANCE : voir Exponentiation

PUISSANCE (Fonction) : voir Analyse et les pages 311, 125.

PYRAMIDE : voir Polyèdres

Q R

QUADRANGLE

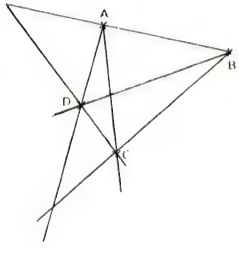


Figure plane formée par quatre points (les sommets) dont trois quelconques ne sont pas alignés et les six droites qui les joignent deux par deux.

QUADRANGLE HARMONIQUE : voir Birapport

QUADRILATÈRE (Quadrilateral) : voir Polygones

QUADRIQUE

On appelle quadrique une surface dont l'équation en coordonnées homogènes (X, Y, Z, T) est du second degré, donc de la forme :

$$aX^2 + a'Y^2 + a''Z^2 + 2bYZ + 2b'ZX + 2b''XY + 2cZT + 2c'XT + 2c''YT + dT^2 = 0.$$

On distingue :

— les quadriques sans point double ou quadriques propres :

- paraboloïdes
- hyperboloïdes
- ellipsoïdes

— les quadriques admettant un seul point double :

- cônes (point double à distance finie)
- cylindres (point double à l'infini)

— les quadriques décomposées en un ensemble de deux plans (sécants, parallèles ou confondus).

QUANTIFICATEUR

Les quantificateurs sont des symboles mathématiques ; ils sont deux :

1. — Le quantificateur existentiel $\exists$, qui se lit : « il existe un », au sens de « il existe au moins un ».

Si E est un ensemble et $\mathcal{R}$ une relation sur cet ensemble :

$$\exists\, x \in E \mid \mathcal{R}$$

se lit : il existe au moins un élément x appartenant à l'ensemble E pour lequel la relation $\mathcal{R}$ soit vraie.

2. — *Le quantificateur universel* $\forall$, qui se lit : « quel que soit » ou « pour tout ». Si E est un ensemble et $\mathcal{R}$ une relation sur cet ensemble :

$$\forall\, x \in \mathrm{E} \mid \mathcal{R}$$

se lit : quel que soit l'élément x de l'ensemble E, la relation $\mathcal{R}$ est vraie, ou : pour tout élément x de l'ensemble E, la relation $\mathcal{R}$ est vraie.

Les deux *propositions* non $(\exists\, x \in \mathrm{E} \mid \mathcal{R})$ et $(\forall\, x \in \mathrm{E} \mid \text{non } \mathcal{R})$ sont équivalentes, d'où un véritable automatisme pour passer d'une proposition à sa négation : on remplace, chaque fois qu'on les rencontre, $\forall$ par $\exists$, $\exists$ par $\forall$ et $\mathcal{R}$ par sa négation. Par exemple, la commutativité d'une loi de composition interne notée par le symbole $\divideontimes$ dans un ensemble E se traduit par :

$$\forall\, a \in \mathrm{E},\ \forall\, b \in \mathrm{E}, \qquad a \divideontimes b = b \divideontimes a.$$

La non-commutativité de la loi $\divideontimes$ dans E se traduira par :

$$\exists\, a \in \mathrm{E},\ \exists\, b \in \mathrm{E}, \qquad a \divideontimes b \neq b \divideontimes a.$$

Dans une formulation mathématique où interviennent les deux quantificateurs, leur ordre d'apparition n'est pas indifférent. Par exemple, la formule

$$\forall\, x \in \mathbb{N},\ \exists\, y \in \mathbb{N}, \quad x \leqslant y$$

signifie : quel que soit l'entier naturel x, il existe un entier naturel y qui soit supérieur ou égal à x.
Si on avait écrit :

$$\exists\, y \in \mathbb{N},\ \forall\, x \in \mathbb{N}, \quad x \leqslant y.$$

(en échangeant l'ordre des deux quantificateurs), cela signifierait qu'il existe au moins un entier naturel y qui soit supérieur à tout entier naturel x.
Les quantificateurs $\forall$ et $\exists$ sont des symboles mathématiques et, comme tels, ne doivent être employés que dans des formules ; ils ne doivent pas être considérés comme des abréviations commodes.

QUARTILE : voir l'article « les Statistiques », page 464.

QUATERNION : voir l'article « les Nombres », page 374 et la page 330.

QUOTIENT : voir Division et la page 281.

RACINE pages 99, 286, 374.

Racine d'une équation

Le mot « racine » est ici synonyme du mot « solution ».
Si on considère l'équation $x^3 - 3x + 2 = 0$, on constate que l'on peut écrire cette équation sous la forme $(x - 1)^2 (x + 2) = 0$. Les racines sont donc $x = 1$ et $x = -2$. La racine 1 est dite racine double, car il y a $(x - 1)$ à la puissance 2 en facteur ; la racine -2 est simple, car $(x + 2)$ ne figure qu'à la puissance 1.
Si, dans une équation algébrique $f(x) = 0$, on peut mettre $(x - a)^n$ en facteur dans $f(x)$ (sans toutefois pouvoir y mettre $(x - a)^{n+1}$), la valeur a de x est une racine de l'équation $f(x) = 0$; on dit qu'elle est racine multiple d'ordre n (simple pour $n = 1$, double pour $n = 2$, triple pour $n = 3$).

Racine d'un polynôme

Si P est un polynôme à coefficients réels (ou complexes), un nombre réel (ou complexe) a est racine du polynôme P si la valeur de $P(x)$ pour $x = a$ est égale à 0 ; autrement dit, a est racine du polynôme P si a est solution de l'équation $P(x) = 0$. Ainsi, le nombre réel 1 est racine du polynôme P, tel que :

$$P(x) = x^2 + 2x - 3 \text{ (en effet, } P(1) = 1 + 2 - 3 = 0).$$

Si a est une racine simple ou multiple d'ordre n de l'équation $P(x) = 0$, on dit que a est racine simple ou multiple d'ordre n du polynôme P.

Un polynôme de degré n à coefficients réels a, au maximum, n racines réelles, tandis qu'un polynôme de degré n à coefficients réels (ou complexes) admet toujours exactement n racines complexes (certaines d'entre elles pouvant être réelles), à condition de compter une racine multiple d'ordre p comme p racines égales. Ainsi, le polynôme $x^2 + x + 1$ n'a aucune racine réelle, mais il a deux racines complexes : $\frac{1}{2}(-1 + i\sqrt{3})$ et $\frac{1}{2}(-1 - i\sqrt{3})$.

Racine carrée d'un nombre réel

Étant donné le nombre réel 4, il existe deux nombres réels (2 et — 2), qui, élevés au carré, sont égaux à 4. Ces deux nombres + 2 et — 2 sont appelés les racines carrées de 4.

De façon générale, étant donné un nombre réel A, on appelle racine carrée de A tout nombre réel a dont le carré soit égal à A ($a^2 = A$).

— Si le nombre A est négatif, il n'y a aucun nombre a qui convienne (car le carré d'un nombre réel est toujours positif).

— Si le nombre A est nul, seul le nombre 0 convient.

— Si le nombre A est positif, il y a toujours deux nombres réels opposés qui conviennent. Par convention, on note $\sqrt{A}$ celui des deux qui est positif (et on l'appelle racine carrée arithmétique de A) ; l'autre est donc égal à $-\sqrt{A}$ (ci-dessus, nous avons vu qu'il y avait deux nombres réels dont le carré est égal à 4 : + 2 et — 2 ; par définition, $\sqrt{4} = +2$).

Racine n-ième d'un nombre réel ◆ ($n \in \mathbb{N}^*$)

On appelle racine n-ième d'un nombre réel A tout nombre réel a qui, élevé à la puissance n, est égal à A ($a^n = A$). Ainsi :

+ 8 a une seule racine cubique (troisième) : + 2 ;
+ 16 a deux racines quatrièmes : + 2 et — 2 ;
— 27 a une seule racine cubique : — 3.

- Si n est impair, tout nombre réel A admet une racine n-ième et une seule.
- Si n est pair : si A est négatif, A n'admet aucune racine n-ième ; si A est positif ou nul, A admet deux racines n-ièmes opposées.

On ne devrait employer le symbole $\sqrt[n]{\ }$ (racine n-ième) que dans le cas précis où le nombre A est positif et où la racine n-ième, que l'on note $\sqrt[n]{A}$, est également positive ($\sqrt[3]{8} = +2$; $\sqrt[4]{16} = +2$).

Il arrive toutefois que l'on trouve des écritures abusives, du genre $\sqrt[3]{-8} = -2$.

Calcul d'une racine n-ième

Il existe un procédé permettant de calculer (on dit extraire) la racine carrée d'un nombre positif◆ ; il n'est pas simple. Il est plus commode (et souvent d'une

◆ La définition donnée ici est encore valable pour les nombres complexes. Toutefois, pour ces nombres, il n'existe pas de nombres positifs ni de nombres négatifs, donc toute la suite n'est pas valable pour eux et il n'est pas question d'employer pour eux le symbole $\sqrt[n]{\ }$.

Il faut remplacer la suite de la définition par le théorème suivant : tout nombre complexe z admet n racines n-ièmes complexes (dont certaines peuvent être réelles). Ainsi, 1 (considéré comme complexe) admet trois racines cubiques : 1, $-\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$ (ces nombres sont notés $1, j, j^2$).

◆ Voir Algorithme.

précision bien suffisante) d'utiliser les logarithmes décimaux (le logarithme de la racine n-ième du nombre x est obtenu en divisant par n le logarithme de x).

RACINE CARRÉE (Fonction) :

voir Fonction et la page 123.

RADIAN : voir Angle

RAISON (d'une suite) ◆

◆ Voir l'article « l'Analyse », pages 101 et 102.

— arithmétique : différence entre deux termes consécutifs ;

— géométrique : quotient de deux termes consécutifs.

RANG ◆

◆ Pour le lecteur peu familiarisé avec les notions d'algèbre linéaire, il est conseillé de lire le début de l'article « l'Algèbre linéaire », pages 91 et suivantes.

Rang d'un système de vecteurs

Étant donné n vecteurs d'un espace vectoriel E, on appelle rang de ce système de vecteurs le plus grand nombre entier p (inférieur ou égal à n) tel qu'un système de p de ces vecteurs soit un système libre, alors qu'un système de $p + 1$ de ces vecteurs serait un système lié. Ainsi, dans l'ensemble des vecteurs du plan, les trois vecteurs : $\vec{V} = \vec{i} + \vec{j}$ $\vec{W} = 2\vec{i} + 3\vec{j}$ $\vec{U} = \vec{i} + 2\vec{j}$ constituent un système de rang 2, car deux quelconques de ces vecteurs constituent un système libre, alors que les trois constituent un système lié $\left(\text{car } \vec{U} + \vec{V} - \vec{W} = \vec{0}\right)$.

Rang d'une matrice carrée d'ordre n ◆

◆ Lire d'abord « l'Algèbre linéaire », pages 89 et suivantes.

Le rang d'une matrice carrée M d'ordre n est le plus grand entier p (inférieur ou égal à n) tel qu'il existe une matrice carrée d'ordre p extraite de la matrice M, qui soit régulière, alors que toute matrice carrée d'ordre $p + 1$, extraite de la matrice, n'est pas régulière.

Ainsi, la matrice $M = \begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix}$ est de rang 2, car la matrice elle-même n'est pas régulière (son déterminant est nul), mais une matrice extraite de M, est d'ordre 2, par exemple $\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$, est régulière (son déterminant vaut $1 \times 1 - 2 \times (-1) = 3$).

RATIONNEL (Nombre) pages 29, 250, 370 et 371, 373, 383, 384.

Un nombre x est un nombre rationnel s'il existe au moins un nombre entier relatif (positif ou négatif) non nul p tel que le produit de p par x soit un nombre entier relatif. Par exemple, 3,6 est un nombre rationnel puisque en le multipliant par 5 on obtient 18, par 10 on obtient 36, etc. Si on le multiplie, par exemple, par 2, on obtient 7,2 qui, lui-même multiplié par 5, donne un entier ; 3,6 et 7,2 sont des nombres rationnels. L'ensemble de ces nombres est noté $\mathbb{Q}$. $\mathbb{Q}$ contient des éléments positifs dont l'ensemble est noté $\mathbb{Q}^+$ et des éléments négatifs dont l'ensemble est noté $\mathbb{Q}^-$.

Tous les nombres entiers relatifs positifs et négatifs, dont l'ensemble est noté $\mathbb{Z}$, sont rationnels ; $\mathbb{Z}$ est inclus dans $\mathbb{Q}$: $\mathbb{Z} \subset \mathbb{Q}$.

Le développement décimal d'un nombre rationnel peut être fini (c'est le cas des nombres décimaux) ou illimité, périodique (et c'est le cas des non-décimaux).

Exemple :

$\frac{19}{28} = 0{,}67\,\overline{8\,571\,42}\,\underline{8\,571\,42}\,\overline{8\,571\,42}\ldots$: non décimal.

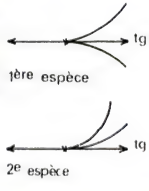
$\frac{23}{20} = 1{,}15$: décimal.

RAYON (d'un cercle) : voir Cercle, Sphère, Disque, Boule

REBROUSSEMENT (Point de) ◆ page 236.

◆ Il est utile de lire d'abord le mot Double (point double d'une courbe).

Un point de rebroussement d'une courbe est un point double de cette courbe pour lequel les tangentes sont confondues. Il y a deux sortes de points de rebroussement : ceux de première espèce, pour lesquels la courbe, dans un voisinage du point de rebroussement, est de part et d'autre de la tangente, et ceux de seconde espèce, pour lesquels la courbe, dans un voisinage du point, est du même côté de la tangente.

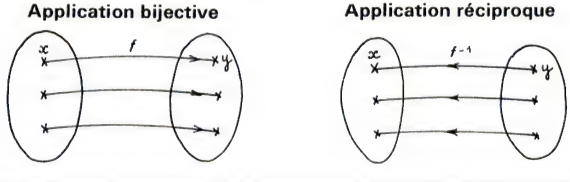


Points de rebroussement

RÉCIPROQUE (Application) ◆ pages 79 et 80, 115, 122, 276, 278, 335.

◆ Lire d'abord Application.

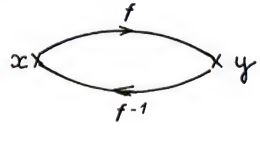
Si f est une application bijective d'un ensemble E vers un ensemble F, chaque élément y de F est l'image par f d'un élément x de E (et d'un seul). On peut donc associer à chaque élément y de F l'élément x de E dont il est image par f, définissant ainsi une application de F sur E, qui est appelée application réciproque de l'application f, et notée f^{-1} (on lit : f moins 1). Cette application est aussi une bijection.



Exemple :

$E = \{1, 2, 3\}$ $\quad F = \{1, 4, 9\}$

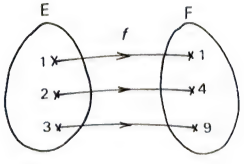
L'application f de E dans F qui à chaque élément de E fait correspondre son carré est une bijection de E sur F. A chaque élément y de F, on peut faire correspondre l'élément x de E dont il est le carré. x est appelé la racine carrée arithmétique de y et noté $\sqrt{y}$. Ici, $y = x^2 \iff x = \sqrt{y}$.



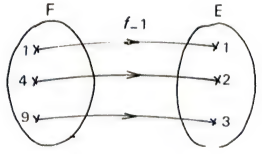
● Dire que l'image par f de x est y est équivalent à dire que l'image par f^{-1} de y est x :

$y = f(x) \iff x = f^{-1}(y)$

Application de E vers F :
« a pour carré »



Application de F vers E :
« a pour racine »



● Si on cherche l'image par f de x ($x \in E$) : y, puis l'image par f^{-1} de y, on retrouve x ; donc l'application composée $f^{-1} \circ f$ ◆ est telle que tout x de E est sa propre image par $f^{-1} \circ f$; c'est donc l'application identique de E :

$$f^{-1} \circ f = I_E$$

◆ Voir Composition des applications.

● Si on cherche l'image de y ($y \in F$) par f^{-1} : x, puis l'image de x par f, on trouve y, donc l'application composée $f \circ f^{-1}$ est telle que tout élément y de F est sa propre image par $f \circ f^{-1}$; l'application $f \circ f^{-1}$ est donc l'application identique de F :

$$f \circ f^{-1} = I_F$$

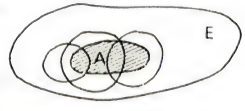
Image réciproque◆

◆ Voir Image.

L'image réciproque d'un sous-ensemble par une application f existe, même si l'application n'admet pas d'application réciproque ; dans le cas où l'application f admet une application réciproque, l'image réciproque d'un sous-ensemble est l'image de ce sous-ensemble par l'application réciproque.

RECOUVREMENT

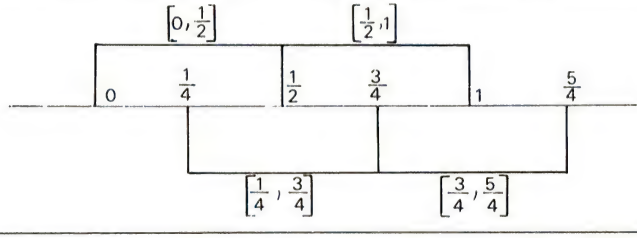
Soient E un ensemble et A une partie de E ; un recouvrement de A est une famille (finie ou infinie) de parties de E dont la réunion contient A.



Recouvrement de A

Par exemple, on peut recouvrir l'intervalle [0, 1] de $\mathbb{R}$, ensemble des nombres réels, par la famille d'intervalles :

$$\left[0, \frac{1}{2}\right], \left[\frac{1}{4}, \frac{3}{4}\right], \left[\frac{1}{2}, 1\right], \left[\frac{3}{4}, \frac{5}{4}\right].$$



Si E est un espace topologique et si le recouvrement de A est formé d'ouverts de E, on dit que ce recouvrement est un recouvrement ouvert.

RECTANGLE : voir Polygones

RÉCURRENCE (Raisonnement par) pages 378, 379.

Supposons qu'une propriété P dépende d'un nombre entier n (par exemple : la somme des n premiers nombres entiers naturels ◆, non nuls, est égale à $\frac{n(n+1)}{2}$), si on démontre les deux propositions suivantes :

1) cette propriété est vraie pour un certain entier n_0,

2) si elle est vraie pour un entier quelconque p, alors elle est vraie pour le suivant $p + 1$.

Cette propriété est vraie pour tous les entiers supérieurs ou égaux à n_0 ◆.

◆ Cette somme est : $1 + 2 + 3 + \ldots + n$.

◆ En général, n_0 est égal à 1.

Par exemple, démontrons que la somme des n premiers nombres entiers est égale à $\frac{n(n+1)}{2}$.

1) Cette propriété est vraie pour $n = 2$; les deux premiers entiers sont 1 et 2, leur somme est égale à 3 ; si on remplace n par 2 dans $\frac{n(n+1)}{2}$, on trouve $\frac{2 \times 3}{2} = 3$.

2) Supposons cette propriété vraie pour p et démontrons qu'alors elle est vraie pour $p + 1$.

Par hypothèse, la somme des p premiers entiers est $\frac{p(p+1)}{2}$. Pour avoir la somme des $p + 1$ premiers entiers, il suffit d'ajouter $p + 1$. On obtient $\frac{p(p+1)}{2} + p + 1$, soit :

$$\frac{p(p+1) + 2(p+1)}{2} = \frac{(p+1)(p+2)}{2}.$$

Donc, la somme des $p + 1$ premiers entiers est égale à $\frac{(p+1)(p+2)}{2}$, ce qui n'est autre que l'expression $\frac{p(p+1)}{2}$, dans laquelle on a remplacé p par $p + 1$.

La formule $1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$ est donc vraie pour tout entier n supérieur ou égal à 2 ◆.

◆ On peut dire qu'elle est aussi vraie pour 1. La somme se réduit au seul nombre 1 et $\frac{n(n+1)}{2}$, dans lequel on remplace n par 1, vaut bien $\frac{1 \times 2}{2} = 1$.

RÉEL (Nombre) ◆ pages 53, 60, 101 à 114, 246, 286, 372 et 373, 385.

Nombre rationnel ou irrationnel. L'ensemble des nombres réels est noté $\mathbb{R}$; il contient des éléments positifs et des éléments négatifs.

L'ensemble des nombres réels positifs est noté $\mathbb{R}^+$, celui des nombres réels négatifs est noté $\mathbb{R}^-$.

Les nombres entiers, les nombres rationnels appartiennent à $\mathbb{R}$.

◆ Voir aussi Coupure.

RÉFÉRENTIEL

Si on ne s'intéresse qu'aux parties d'un ensemble E fixé, cet ensemble E est qualifié de référentiel.

RÉFLEXIVE◆ pages 280 et 281, 298, 299.

◆ Lire d'abord Relation.

Soit E un ensemble et $\mathcal{R}$ une relation binaire définie dans cet ensemble ; cette relation $\mathcal{R}$ est réflexive si, quel que soit l'élément x de E, cet élément est en relation avec lui-même par la relation $\mathcal{R}$:

$$\forall x \in E, \quad x \mathcal{R} x$$

(lire : quel que soit x appartenant à E, x est en relation avec x).

Par exemple, dans $\{1, 2, 3\}$, la relation « divise » est réflexive, car 1 divise 1, 2 divise 2 et 3 divise 3 ; par contre, dans $\{0, 1, 2\}$, la relation « divise » n'est pas réflexive, car 0 ne divise pas 0.

Diagramme cartésien d'une relation réflexive

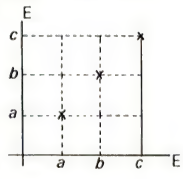
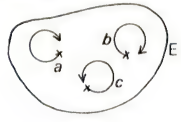


Diagramme sagittal



REGULIER

Elément régulier : voir l'article « l'Algèbre », page 25.

Matrice régulière : voir Matrice.

Polygone régulier : voir Polygone.

Pour les diagrammes sagittal et cartésien, se reporter à l'article « les Ensembles », pages 274, 275 et à Diagramme.

RELATIF (Nombre) : voir Entier (Nombre) et les pages 29, 46, 369 et 370, 380 à 383.

RELATION BINAIRE pages 155, 173, 191, 242, 252, 273 à 285, 276, 298, 299, 314, 319, 334, 360, 413, 447.

On peut dire qu'on a défini une relation binaire (on l'appelle ainsi parce qu'elle met en jeu deux éléments) entre les éléments d'un ensemble E (appelé ensemble de départ ou source) et les éléments d'un ensemble F (ensemble d'arrivée ou but) si l'on a défini une propriété telle que, pour tout couple (x, y), x appartenant à E, y appartenant à F, ou bien ce couple vérifie la propriété et on dit alors que x est en relation avec y ($x \mathcal{R} y$) ou ce couple ne vérifie pas la propriété, et on dit que x n'est pas en relation avec y.

Exemple

Soient deux ensembles $E = \{1, 2, 3\}$ et $F = \{1, 2, 3, 4, 5\}$. L'élément x de E est dit en relation « divise » avec y de F si x divise y.

Les couples vérifiant la relation sont les suivants :

(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 2), (2, 4), (3, 3)

Les couples vérifiant la relation constituent le graphe ◆ de la relation. Une relation est connue si on connaît E, F et son graphe G. Se donner une relation, c'est se donner deux ensembles E et F et un sous-ensemble G du produit cartésien $E \times F$. Un couple (x, y) de $E \times F$ vérifie la relation $\mathcal{R}$ si et seulement si $(x, y) \in G$. Deux relations $\mathcal{R} = (E, F, G)$ et $\mathcal{R}' = (E', F', G')$ sont égales si et seulement si $E = E'$, $F = F'$, $G = G'$.

Parmi toutes les relations binaires, sont très importantes : les relations fonctionnelles, les relations d'ordre, les relations d'équivalence.

◆ Voir l'article « les Ensembles », pages 274 et 275.

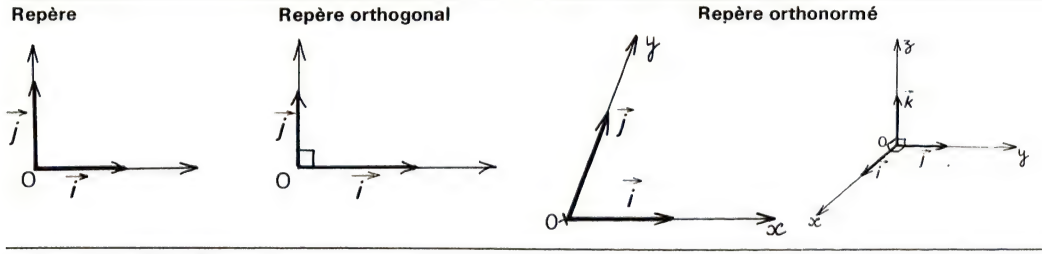
RÉPARTITION : voir l'article « les Statistiques », pages 459, 475.

REPÈRE

Dans le plan, un repère cartésien affine est constitué d'un point O (appelé origine du repère) et de deux vecteurs $\vec{i}$ et $\vec{j}$ non colinéaires. On note le repère $(O, \vec{i}, \vec{j})$

(on a souvent l'habitude de considérer les deux axes Ox et Oy respectivement portés par $\vec{i}$ et par $\vec{j}$ et de même sens et on parle aussi de repère (Ox, Oy)).

Si les vecteurs $\vec{i}$ et $\vec{j}$ sont orthogonaux, le repère est qualifié d'orthogonal.



Si les vecteurs $\vec{i}$ et $\vec{j}$ sont orthogonaux et ont même longueur, le repère est qualifié d'orthonormé.

De la même façon, dans l'espace, un repère cartésien affine est constitué d'un point O et de trois vecteurs non coplanaires $\vec{i}$, $\vec{j}$, $\vec{k}$.

Si les trois vecteurs $\vec{i}$, $\vec{j}$, $\vec{k}$ sont deux à deux orthogonaux, le repère est qualifié d'orthogonal.

Si $\vec{i}$, $\vec{j}$, $\vec{k}$ sont deux à deux orthogonaux et ont même longueur, le repère est qualifié d'orthonormé.

REPRÉSENTANT

Soient E un ensemble sur lequel est définie une relation d'équivalence $\mathcal{R}$, et x un élément de E ; dans la classe d'équivalence de x, il y a tous les éléments de E équivalents à x (et eux seulement) ; si y est un élément de cette classe, la classe d'équivalence de y est égale à la classe d'équivalence de x, donc la donnée de l'un quelconque des éléments de cette classe suffit à la définir. Chaque élément de la classe d'équivalence est appelé un représentant de cette classe (en particulier, x est un représentant de $\mathrm{Cl}(x)$).

RESTE : voir Division

RESTRICTION ◆

◆ Lire d'abord Application.

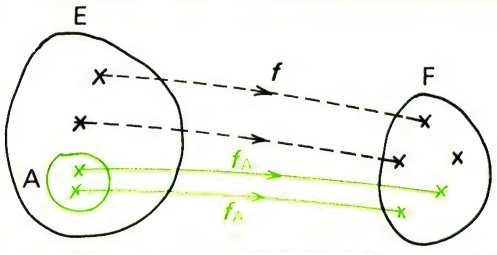
Si f est une application d'un ensemble E dans un ensemble F et si A est une partie non vide de E, l'application de A dans F qui à tout x de A fait correspondre $f(x)$:

$$f : x \longmapsto f(x)$$
$$x \in \mathrm{E} \qquad f(x) \in \mathrm{F}$$

s'appelle restriction de f à A et on la note f_{A} ou f/A :

$$f_{\mathrm{A}} : x \longmapsto f(x)$$
$$x \in \mathrm{A} \qquad f(x) \in \mathrm{F}$$

Cette restriction de f à A est unique.



f_A est la restriction de l'application f.

RÉTICULE (Ensemble) : voir Treillis

RÉUNION pages 22, 23, 25, 26, 190, 249, 265, 363, 364.

Si E et F sont deux ensembles, on appelle réunion de E et de F et on note E ∪ F (lire : E union F) l'ensemble des éléments qui appartiennent au moins à l'un des deux ensembles :

$$E \cup F = \{ x \mid x \in E \text{ ou } x \in F \}$$

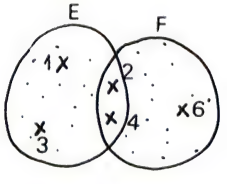
(lire : E union F est égal à l'ensemble des x tels que x appartienne à E ou x appartienne à F).

Quels que soient les deux ensembles E et F, la réunion de E et de F est égale à la réunion de F et de E :

$$E \cup F = F \cup E$$

Quel que soit l'ensemble E, la réunion de E avec l'ensemble vide (noté ∅) est égale à E :

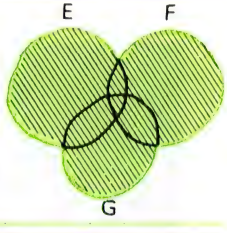
$$E \cup \varnothing = E$$



$E = \{1, 2, 3, 4\}$

$F = \{2, 4, 6\}$

$E \cup F = \{1, 2, 3, 4, 6\}$



E ∪ F ∪ G

Si on introduit un troisième ensemble G, quels que soient les trois ensembles E, F et G, la réunion de E avec la réunion de F et G est égale à la réunion de la réunion de E et de F avec G :

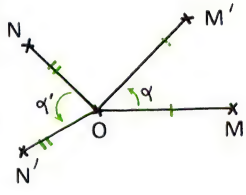
$$E \cup (F \cup G) = (E \cup F) \cup G$$

Cet ensemble est appelé réunion des trois ensembles E, F et G et noté sans parenthèses : E ∪ F ∪ G.

Si $E_1, E_2, \ldots, E_n$ sont n ensembles, on note $\bigcup_{i=1}^{n} E_i$ la réunion de ces ensembles.

RÉVOLUTION (Surface de) : voir Surface de révolution

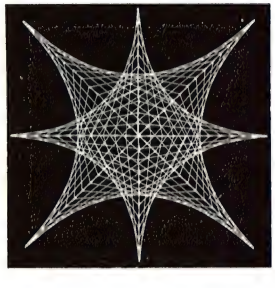
ROTATION pages 77, 154, 453.



Dans le plan, étant donné un point O et un angle α, on appelle rotation de centre O et d'angle α la transformation qui à chaque point M du plan fait correspondre le point M' tel que la longueur de OM' soit égale à la longueur de OM et que l'angle du vecteur $\overrightarrow{OM}$ avec le vecteur $\overrightarrow{OM'}$ soit égal à l'angle α.
Si l'angle α est l'angle nul, la rotation est la transformation identique (car tout point M se transforme alors en lui-même).
La composée de deux rotations de même centre O et d'angles respectifs α et α' est la rotation de centre O et d'angle $\alpha + \alpha'$.
Muni de la loi de composition des applications, l'ensemble des rotations de centre O a structure de groupe commutatif.

Actuellement, on préfère suivre le chemin inverse et définir les angles à partir des rotations. On définit alors une rotation comme étant une application linéaire de l'espace vectoriel $\mathbb{R}^2$ dans lui-même dont la matrice, dans une base orthonormée, est de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, où a et b sont deux nombres réels dont la somme des carrés est égale à 1 ($a^2 + b^2 = 1$).
Si le plan est orienté, l'angle α de la rotation est tel que :

$$\begin{cases} \cos \alpha = a \\ \sin \alpha = b. \end{cases}$$



S

SAGITTAL (Diagramme) : voir Diagramme et les pages 275, 279 et 280.

SCALAIRE pages 60, 61, 65, 72.

En algèbre linéaire : voir l'article « l'Algèbre linéaire», page 60.

Matrice scalaire

Une matrice scalaire est une matrice carrée dont tous les éléments situés sur la diagonale principale sont égaux et dont tous les autres éléments sont nuls. Une telle matrice commute avec toute autre matrice carrée de même ordre qu'elle, c'est-à-dire que, si M est une matrice scalaire d'ordre n, quelle que soit la matrice carrée N d'ordre n, on a : $N \times M = M \times N$.

Exemple : matrice scalaire d'ordre 3
$M = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$.

Produit scalaire : voir Produit.

SECTION COMMENÇANTE◆

◆ Lire d'abord Ordre, Ordonné, Relation.

Si E est un ensemble ordonné (par la relation d'ordre notée $\leqslant$), on dit qu'une partie S de E est une section commençante si tout élément de E, qui est inférieur à au moins un élément de S, appartient à S.
Une section commençante de E est ouverte si elle est non vide et n'admet pas de plus grand élément ; par exemple, si a est un élément de E, l'ensemble des x de E tels que x soit inférieur strictement à a est une section commençante ouverte. Une section commençante de E est fermée si elle admet un plus grand élément ; par exemple, si a est un élément de E, l'ensemble des x de E, inférieurs ou égaux à a, est une section commençante fermée.

SEMI-OUVERT : voir Intervalle

SÉPARÉ◆

◆ Lire d'abord Topologie.

Un espace topologique E est séparé si, quels que soient les points distincts x et y de E, on peut trouver un voisinage de V de x et un voisinage W de y n'ayant aucun point commun.
$\mathbb{R}$, ensemble des nombres réels, est un espace séparé, car, si x et y sont deux nombres réels distincts, on peut toujours trouver un intervalle entourant x et un intervalle entourant y qui n'aient aucun point commun.

SÉRIE pages 162, 235, 328.

Étant donné une suite ◆ de nombres réels $u_0, u_1, u_2, \ldots, u_n, \ldots$, on appelle série de terme général u_n la suite (S_n) définie par : $S_0 = u_0$, $S_1 = u_0 + u_1$, $S_2 = u_0 + u_1 + u_2$, $\ldots$, $S_n = u_0 + u_1 + u_2 + \ldots + u_n$. S_n est la somme des n premiers termes de la suite.

◆ Voir l'article « l'Analyse », pages 107 et suivantes.

Ainsi, étant donné la suite géométrique ◆ du premier terme 1 et de raison $\frac{1}{2}$, c'est-à-dire la suite : $u_0 = 1$, $u_1 = \frac{1}{2}$, $u_3 = \frac{1}{8}$, $\ldots$, $u_n = \frac{1}{2^n}$, on peut écrire la « série géométrique » : $S_0 = 1$, $S_1 = 1 + \frac{1}{2}$, $S_2 = 1 + \frac{1}{2} + \frac{1}{4}$, $S_3 = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8}$, $\ldots$, $S_n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{2^n}$, $\ldots$

◆ Voir l'article « l'Analyse », page 111.

Étant donné la suite harmonique ◆ $u_0 = 1$, $u_1 = \frac{1}{2}$, $u_2 = \frac{1}{3}$, $\ldots$, $u_n = \frac{1}{n+1}$, on peut définir la « série harmonique » $S_1 = 1$, $S_2 = 1 + \frac{1}{2}$, $S_3 = 1 + \frac{1}{2} + \frac{1}{3}$, $S_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}$, $\ldots$, $S_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n}$.

◆ Voir l'article « l'Analyse », page 107.

Série convergente ; série divergente

On dit que la série converge s'il existe un nombre fini S ◆ (appelé somme de la série) tel que S_n tende vers S quand n tend vers $+\infty$. Dans le cas où la série ne converge pas, on dit qu'elle diverge.

◆ On note alors : $S = \sum_{n=0}^{+\infty} u_n$.

Exemples

— La série géométrique, de terme général ◆ $u_n = q^n$, est convergente si q est compris entre 0 et 1 car :

$$S_n = 1 + q + q^2 + \ldots + q^n = \frac{1 - q^{n+1}}{1 - q} ;$$

quand n tend vers $+\infty$, q^{n+1} tend vers 0 si q est inférieur à 1 et S_n a pour limite $\frac{1}{1-q}$.

◆ $u_0 = 1$, $u_1 = q$, $u_2 = q^2$, $\ldots$ donc : $S_0 = 1$, $S_1 = 1 + q$, $S_2 = 1 + q + q^2$, $\ldots$

— La série de terme général $u_n = \frac{1}{n!}$ ◆ est convergente. Sa somme est le nombre e ◆.

◆ $n!$ = factorielle n.

◆ e est la base du logarithme népérien (voir ce mot).

— La série harmonique, de terme général $u_n = \frac{1}{n+1}$ ◆, est divergente $\left(\text{car } 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n} \text{ ne tend pas vers une limite finie quand } n \text{ tend vers } + \infty\right)$.

◆ $u_0 = 1$, $u_1 = \frac{1}{2}$, $u_2 = \frac{1}{3}$, donc : $S_0 = 1$, $S_1 = 1 + \frac{1}{2}$, $S_2 = 1 + \frac{1}{2} + \frac{1}{3}$, $\ldots$

Si une série converge, alors le terme général u_n tend vers 0 quand n tend vers $+\infty$; mais u_n peut tendre vers 0 sans que la série converge $\left(\text{la série harmonique a bien son terme général } u_n = \frac{1}{n+1} \text{ qui tend vers 0, mais elle ne converge pas}\right)$.

On ne change pas la nature d'une série (c'est-à-dire le fait qu'elle soit convergente ou divergente) en supprimant un nombre fini de termes de cette série (naturel-

lement, dans le cas d'une série convergente, la somme S sera un peu modifiée par la suppression de ces termes).

Série à termes positifs
Une série est dite à termes positifs si, à partir d'un certain rang, tous les u_n sont positifs ou nuls.

Série alternée
Une série alternée est une série où u_n est alternativement positif et négatif. Pour qu'une telle série converge, il faut et il suffit que la valeur absolue de son terme général tende vers 0, en décroissant, quand n tend vers $+\infty$.

La série de terme général $u_n = \frac{(-1)^n}{n+1}$ est une série alternée convergente.

Série absolument convergente ; série semi-convergente
Si la série de terme général égal à la valeur absolue de u_n : $|u_n|$ converge, on dit que la série u_n est absolument convergente.
Une série absolument convergente est convergente ; mais une série peut être convergente sans être absolument convergente ; on dit alors qu'elle est semi-convergente (ainsi, la série $u_n = \frac{(-1)^n}{n+1}$ est une série convergente, mais pas absolument convergente, car $|u_n| = \frac{1}{n+1}$, et la série de terme général $\frac{1}{n+1}$ est divergente).

Série entière
Une série entière est une série dont le terme général u_n est de la forme : $a_n \cdot x^n$, où x est un nombre réel quelconque.

Série trigonométrique
Série dont le terme général u_n est de la forme : $a_n \cos nx + b_n \sin nx$.

SÉRIE CHRONOLOGIQUE (ou chronique) page 235.

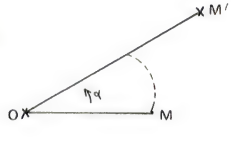
Suite de valeurs prises par une variable au cours du temps. Par exemple : production annuelle de blé en France de 1964 à 1972.

SÉRIE STATISTIQUE : voir l'article « les Statistiques », pages 471 à 484, 486, 502, 510 à 512.

SIGNIFICATION (Seuil de) : voir l'article « les Statistiques », pages 509, 514, 515.

SIMILITUDE

Étant donné, dans le plan, une rotation et une homothétie de même centre O, on appelle similitude la composée de ces deux transformations. Si α est l'angle de la rotation et k le rapport d'homothétie, tout point M du plan est transformé en M′ tel que le vecteur $\overrightarrow{OM'}$ soit égal au produit par k du vecteur $\overrightarrow{OM}$ et que l'angle de $\overrightarrow{OM}$ avec $\overrightarrow{OM'}$ soit égal à α.
Si α est nul et k égal à 1, la similitude est la transformation identique.
La composée de deux similitudes de centre O est une similitude de centre O.



L'ensemble des similitudes de centre O, muni de la loi de composition des applications, est un groupe commutatif ◆.

◆ Voir l'article « l'Algèbre », page 28.

SINGLETON page 269.

Un ensemble constitué par un seul élément est appelé singleton. Si cet élément est désigné par x, le singleton est noté $\{ x \}$. Il ne faut pas confondre l'élément x avec le singleton $\{ x \}$ (le premier est élément du second : $x \in \{ x \}$, x est le seul élément de $\{ x \}$). Par exemple, à partir de l'ensemble $\varnothing$ (vide), on peut former le singleton $\{ \varnothing \}$; le premier est un ensemble qui n'a aucun élément, tandis que le second est un ensemble qui a un élément : l'ensemble vide.

SINGULIER

Un point A d'une surface est dit point singulier de cette surface s'il n'existe pas de plan tangent en A à la surface. Par exemple, le sommet d'un cône est un point singulier pour le cône.

Pour intégrale singulière d'une équation différentielle : voir Équation.

SINUS (Fonction) : voir Trigonométriques (Fonctions)

SOMME : voir Addition

Somme de Riemann : voir l'article « l'Analyse », pages 136, 137, 153.

SOMMET : voir Cône, Polyèdres, Polygones, Pyramide

SONDAGE : voir l'article « les Statistiques », pages 457, 507 à 518.

SOUS-ANNEAU : voir l'article « l'Algèbre », page 49.

SOUS-GROUPE : voir l'article « l'Algèbre », pages 37 à 39.

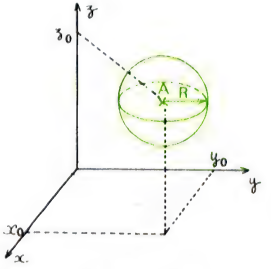
SPEARMAN (Coefficient de) : voir l'article « les Statistiques », page 484.

SPHÈRE page 328.

En géométrie dans l'espace, on appelle sphère l'ensemble des points M de l'espace situés à une même distance R (R nombre réel positif) d'un point A. Le point A est appelé centre de la sphère et le nombre R, le rayon de la sphère.

Si l'espace est muni d'un repère orthonormé Ox, Oy, Oz, l'équation de la sphère de centre A, de coordonnées (x_0, y_0, z_0) et de rayon R est :

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2$$



Surface de la sphère : $4\pi R^2$.

Volume de la sphère : $\frac{4}{3}\pi R^3$.

Sphère dans un espace métrique quelconque

Soit E un ensemble muni d'une distance ◆ d ; x est un point de E et r un nombre réel positif.

◆ Lire ce mot.

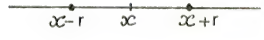
On appelle sphère de centre x et de rayon r et on note $S(x, r)$ l'ensemble des éléments y de E dont la distance $d(x, y)$ à x est égale à r :

$$S(x, r) = \{ y, y \in E \mid d(x, y) = r \}.$$

Par exemple, l'ensemble $\mathbb{R}$ des nombres réels peut être muni d'une distance d (la distance $d(x, y)$ de deux nombres réels étant la différence entre le plus grand et le plus petit) ; la sphère de centre x et de rayon r est formée des nombres réels y dont la distance à x est égale à r ; il y en a deux :

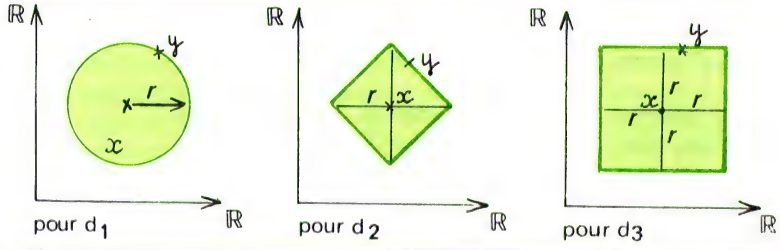
$x + r$ (pour lequel la différence $x + r - x = r$)
$x - r$ (pour lequel la différence $x - (x - r) = r$)

$$S(x, r) = \{ x - r, x + r \}.$$



L'ensemble $\mathbb{R}^2$ des couples (x, y) pour lesquels x et y sont réels peut être muni d'une distance de trois façons différentes ◆. Pour chacune d'elles, voici la forme de la sphère de centre x et de rayon r :

◆ Voir Distance.



STABLE ◆ pages 35 à 37.

◆ Il est conseillé de lire d'abord Application et Image.

Partie stable pour une application

Soient E un ensemble, f une application de E dans E et A une partie de E ; A est stable pour l'application f si l'image $f(A)$ est contenue dans A (c'est-à-dire si, pour tout x de A, l'image $f(x)$ appartient à A).

Exemple : la partie [0, 1] (intervalle fermé) de $\mathbb{R}$ (ensemble des nombres réels) est stable pour l'application $f : x \longmapsto x^2$, car, pour tout x de [0, 1], $x^2 \in [0, 1]$.

Partie stable pour une loi de composition interne ◆

◆ Il est conseillé de lire d'abord l'article « l'Algèbre », pages 17 et suivantes.

Soit E un ensemble muni d'une loi de composition interne notée ✳ ; une partie A de E est stable pour la loi ✳ si, pour tout couple (x, y) d'éléments de A, le composé x ✳ y est élément de A.

Exemple : dans $\mathbb{Z}$, l'ensemble des entiers pairs est une partie stable pour l'addition car la somme de deux entiers pairs est un entier pair ; par contre, l'ensemble des entiers impairs n'est pas stable pour l'addition, car la somme de deux entiers impairs est un entier pair (et non impair).

Partie stable pour une loi de composition externe

Si E est un ensemble, muni d'une loi de composition externe notée . (d'ensemble d'opérateurs Ω), une partie A de E est stable pour cette loi si, pour tout λ de Ω et pour tout x de A, le composé $\lambda . x$ est élément de A.

STATISTIQUES : voir l'article pages suivantes et les pages 163, 167 à 170, 172, 176, 414.

Les statistiques

par Michèle Pascal

Le calcul des probabilités est né au XVII^e siècle. Pascal (1623-1662) et Fermat (1601-1665) l'ont créé. Une des applications de ce calcul est la statistique mathématique. Depuis la fin du XIX^e siècle, les méthodes de la statistique sont employées dans un grand nombre de sciences et de techniques : astronomie, physique, chimie, biologie, linguistique, économie, sociologie, psychologie, médecine, assurance, industrie, etc.

La statistique cherche à dégager les lois caractérisant le comportement de grands ensembles. Ces ensembles étant le siège de phénomènes aléatoires, la théorie des probabilités donne un modèle de comportement de ces phénomènes ◆.

◆ Cette étude se divise en deux grandes parties : la statistique descriptive et la statistique mathématique. Pour la compréhension de cette deuxième partie, il est nécessaire de posséder de solides notions de calcul des probabilités.

LA STATISTIQUE DESCRIPTIVE

La population

La statistique a donc pour objet l'étude des grands ensembles, c'est-à-dire des ensembles contenant un grand nombre d'éléments. On les désigne souvent sous le nom de « population ». Naturellement, ces « populations » ne se limitent pas aux ensembles de personnes. Ce sera, par exemple : l'ensemble des voitures immatriculées à Paris en 1972, l'ensemble des accidents d'avions survenus dans le monde entier pendant les vingt dernières années, etc.

Le caractère

Lorsqu'on a défini avec précision l'ensemble que l'on désire étudier, il faut encore choisir le caractère sous lequel on veut envisager cette population. Par exemple : ensemble des voitures immatriculées à Paris en 1972 étudié sous le caractère « puissance de ces voitures en C.V. » ; ensemble des accidents d'avions survenus depuis vingt ans étudié sous le caractère « nationalité de la compagnie d'aviation », etc. Le caractère peut être quantitatif, c'est-à-dire mesurable (c'est le

cas du premier exemple) ou qualitatif et, par suite, non mesurable (c'est le cas du deuxième). Nous ne nous occuperons que des premiers.

L'enquête

Les deux données précédentes (population et caractère) étant fixées, on mesure alors la valeur du caractère sur chaque élément de la population. On obtient un ensemble de mesures, appelé souvent ensemble d'observations. Cet ensemble est une série statistique. La mesure dont nous venons de parler constitue ce qui est appelé l'« enquête ».
Cette enquête peut être exhaustive, c'est-à-dire qu'elle porte sur tous les éléments de la population étudiée. Elle peut être partielle et ne porter que sur une partie de la population. C'est une enquête par sondage. La série statistique n'est alors l'image que d'une partie de la population, d'un « échantillon ».
Les éléments faisant partie de l'échantillon peuvent être prélevés sur la population totale « au hasard ». Exemple : on veut étudier la durée de vie des ampoules électriques d'une certaine fabrication et, pour cela, on en prélève 500 au hasard. Dans ce cas, le calcul des probabilités pourra être utilisé (ainsi que nous le verrons plus tard) pour inférer, des caractéristiques de l'échantillon, des conclusions relatives à la population entière.
Le sondage n'est pas toujours pratiqué au hasard. Il peut être le résultat d'un choix raisonné (méthode des quotas). C'est souvent le cas dans les sondages d'opinion, dans les études de marché. On fait un échantillon représentatif de l'ensemble en s'aidant des connaissances que l'on peut avoir sur la structure de cet ensemble. Le procédé est commode, mais le calcul des probabilités ne peut être utilisé qu'avec certaines précautions.

Série statistique à caractère continu, à caractère discontinu

L'enquête terminée, le statisticien est en possession d'un énorme ensemble de nombres qu'il doit présenter d'une façon claire. Ce premier travail fait l'objet de la statistique descriptive.
Nous allons en exposer les principales démarches en prenant deux exemples.
Pour faciliter la présentation, nous choisirons une petite population, mais il est essentiel de retenir qu'il ne peut y avoir de résultats valables qu'à partir de l'étude des grands ensembles.

Soit par exemple l'ensemble des propriétés rurales de la région X... étudiées sous le caractère « nombre de personnes actives travaillant à la propriété ».

Le résultat de l'enquête donne la *série statistique suivante* ◆ :

◆ La population comprend 40 propriétés. On mesurera sur chacune d'elles la valeur du caractère, c'est-à-dire le nombre de personnes actives. On obtiendra 40 nombres représentant chacun un nombre de personnes actives.

1	2	1	3	2	1	2	2	
2	2	1	3	3	1	2	2	
1	3	1	5	5	2	2	2	**Série 1**
3	4	2	1	1	2	3	3	
4	4	4	1	1	2	3	2	

Étudions maintenant l'ensemble de ces mêmes propriétés sous le caractère « surface cultivable exprimée en hectares ». Nous obtenons la série statistique suivante :

20	42	25	48	38	35	50	35	
35	40	26	52	48	30	27	36	
22	31	30	70	68	36	32	42	**Série 2**
45	65	35	25	30	38	55	52	
60	70	58	26	26	45	56	30	

Dans chacune de ces séries, le premier travail consiste à ordonner les termes de la série par valeurs croissantes ◆ :

◆ Dans la pratique, ce triage est fait par des machines à cartes perforées, chaque carte représentant un élément de la série.

1	1	1	2	2	2	3	4	
1	1	2	2	2	3	3	4	
1	1	2	2	2	3	3	4	**Série 1**
1	1	2	2	2	3	3	5	
1	1	2	2	2	3	4	5	

20	26	30	35	38	45	52	60	
22	26	30	35	39	45	52	65	
25	27	32	36	40	48	55	68	**Série 2**
25	30	35	36	42	48	56	70	
26	30	35	38	42	50	58	70	

1	2	3	4	5
1	2	3	4	5
1	2	3	4	
1	2	3	4	
1	2	3		
1	2	3		
1	2	3		
1	2	3		
1	2			
1	2			
1	2			
	2			
	2			
	2			**Série 1**
	2			

Puis on fait une partition dans l'ensemble de ces mesures, chaque sous-ensemble obtenu étant appelé classe.
Pour la série 1, il est naturel de mettre dans une même classe tous les éléments égaux.

20	25	30	35	40	50	60
22	25	30	35	42	52	65
	26	30	35	42	52	68
	26	30	35	45	55	70
	26	32	36	45	56	70
	27		36	48	58	
			38	48		
			38			Série 2
			39			

Pour la deuxième série, on groupera dans une même classe les éléments de valeurs voisines. Le choix de l'amplitude d'une classe est arbitraire, mais un nombre trop élevé de classes fait apparaître des irrégularités dues aux effectifs trop faibles par classe, et un nombre trop restreint de classes entraîne une perte d'information.

On peut alors résumer les résultats dans les deux tableaux suivants :

Série 1 : Répartition des propriétés rurales de la région X... d'après le nombre de personnes actives travaillant à la propriété :

personnes actives	1	2	3	4	5	total
propriétés	11	15	8	4	2	40

Série 2 : Répartition des propriétés rurales de la région X... d'après leur surface cultivable exprimée en hectares.

surface	20 à < 25	25 à < 30	30 à < 35	35 à < 40	40 à < 50	50 à < 60	supérieur à 60	total
propriétés	2	6	5	9	7	6	5	40

La première série est une série à caractère discontinu ; tous les éléments d'une même classe ont même valeur. La deuxième est une série à caractère continu ; tous les éléments d'une même classe ont des valeurs voisines comprises entre deux limites. On appelle valeur centrale de la classe la demi-somme des limites.
Sur la deuxième ligne de chaque tableau se trouvent les *effectifs de classe*.

Effectifs cumulés

On peut calculer les effectifs cumulés en additionnant de proche en proche les effectifs de gauche à droite ou de droite à gauche ◆. On obtient ainsi : de gauche à droite, le nombre d'éléments de la série dont la valeur est *au plus* égale à la valeur de la classe correspondante (par exemple, il y a 13 propriétés dont la surface est au plus égale à 35 ha) ; de droite à gauche, le nombre d'éléments de la série dont la valeur est *au moins* égale à la valeur de la classe correspondante; par exemple, l'effectif cumulé 14 — troisième classe, série 1 — indique qu'il y a 14 propriétés qui occupent au moins 3 personnes.

◆ Par exemple, pour la série 2, de gauche à droite ; 2, 2 + 6, 8 + 5, etc. ; de droite à gauche ; 5, 5 + 6, 11 + 7, etc.

Série 1	personnes	1	2	3	4	5
	effectifs	11	26	34	38	40
	cumulés	40	29	14	6	2

Série 2	surfaces	20 à < 25	25 à < 30	30 à< 35	35 à < 40	40 à < 50	50 à < 60	60 à 70
	effectifs	2	8	13	22	29	35	40
	cumulés	40	38	32	27	18	11	5

Fréquence de classe

La fréquence d'une classe est égale au rapport de l'effectif de cette classe à l'effectif total.
La somme des fréquences est égale à 1.

Série 1	personnes	1	2	3	4	5	total
	fréquences	0,275	0,375	0,2	0,1	0,05	1
		$\frac{11}{40}=0{,}275$	$\frac{15}{40}=0{,}375$	$\frac{8}{40}=0{,}2$	$\frac{4}{40}=0{,}1$	$\frac{2}{40}=0{,}05$	

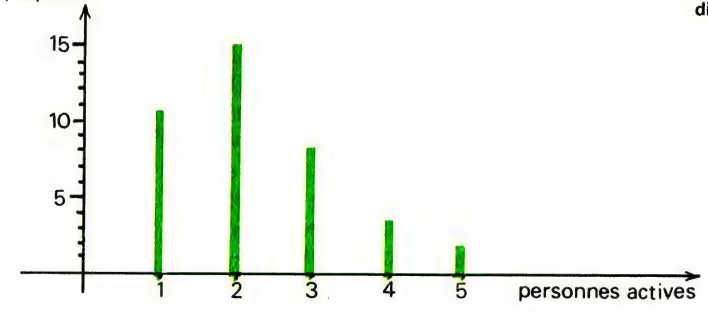
Série 2	surfaces	20 à < 25	25 à < 30	30 à < 35	35 à < 40	40 à < 50	50 à < 60	supérieur à 60	total
	fréquences	0,05	0,15	0,125	0,225	0,175	0,15	0,125	1

Les deux tableaux que nous avons dressés, complétés par les notions d'effectifs cumulés et de fréquences, donnent une vue d'ensemble déjà très satisfaisante de la répartition des propriétés d'après le caractère étudié. Remarquons que, dans la série à caractère continu, le tableau entraîne une perte d'information qui est compensée par une vue plus synthétique de l'ensemble des résultats.

Cet aperçu est encore amélioré dans les représentations graphiques qui traduisent ces résultats d'une façon plus sensible.

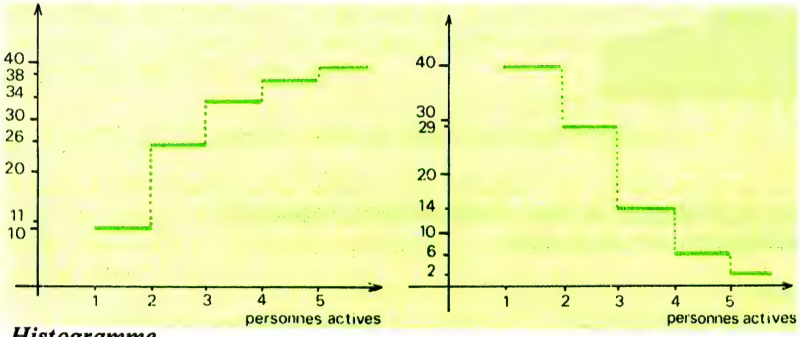
Diagramme en bâtons

série à caractère discontinu



La série à caractère discontinu se représente par un diagramme en bâtons. On porte en abscisse les valeurs du caractère et en ordonnée les effectifs de classe. Des points obtenus, on abaisse des bâtons sur l'axe des abscisses.

Les effectifs cumulés se représentent par un diagramme en escalier. On porte en abscisse les valeurs du caractère et en ordonnée les effectifs cumulés.



Histogramme

La série à caractère continu se représente par un histogramme. L'histogramme est constitué d'une série de rectangles. Les bases des rectangles sont formées par un segment de l'axe des abscisses dont la longueur est proportionnelle à l'intervalle de classe (différence entre les deux limites de la classe).

La surface de chaque rectangle est proportionnelle à l'effectif de classe correspondant.

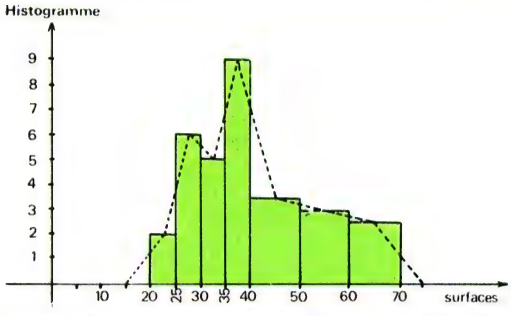
Pour construire l'histogramme, il faut d'abord calculer les hauteurs des rectangles, en divisant l'effectif de classe par l'intervalle de classe. Ces hauteurs seront portées en ordonnée. Les hauteurs des rectangles sont parfois appelées densité de classe.

Calcul des hauteurs

On remarquera qu'on a pris pour unité d'intervalle le plus petit de ces intervalles de classe, les autres étant mesurés avec cette nouvelle unité. Dans notre exemple, le plus petit intervalle étant 5, il est pris comme unité. Un intervalle de 10 se trouve être égal à deux unités.

◆ En indice sont indiquées les valeurs des intervalles, mesurées avec la nouvelle unité.

Série 2	surface	20 à < 25	25 à < 30	30 à < 35	35 à < 40	40 à < 50	50 à < 60	60 à < 70
	propriétés	2	6	5	9	7	6	5
	intervalle de classe	5_1◆	5_1	5_1	5_1	10_2	10_2	10_2
	hauteurs	2	6	5	9	3,5	3	2,5

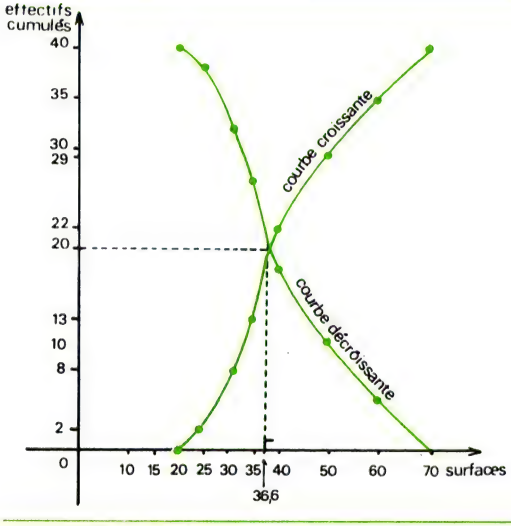


On appelle *polygone de fréquence* la ligne brisée fermée joignant le milieu des bases supérieures des rectangles.

Graphiques cumulatifs

Les effectifs cumulés se représentent par un diagramme cumulatif. Si les effectifs sont cumulés dans le sens des valeurs croissantes du caractère, on porte en abscisse la *limite supérieure* des classes et en ordonnée le nombre cumulé correspondant◆. On joint les points

◆ Voir page 459.



par une courbe continue, le caractère étant assimilé à une variable continue. Si les effectifs sont cumulés dans le sens des valeurs décroissantes du caractère, on porte en abscisse la *limite inférieure* des classes et en ordonnée le nombre cumulé correspondant.
On joint encore les points par une courbe.
L'ordonnée d'un point quelconque de la courbe croissante donne le nombre de fermes dont la surface est au plus égale à son abscisse.
L'ordonnée d'un point quelconque de la courbe décroissante donne le nombre de fermes dont la surface est au moins égale à son abscisse.
Le point de rencontre de ces deux courbes joue un rôle privilégié dont nous parlerons plus tard.

CARACTÉRISTIQUES DES SÉRIES STATISTIQUES

Nous allons maintenant essayer de résumer les séries statistiques par un certain nombre de caractéristiques qui permettent de s'en faire une idée à l'aide de quelques données et de les comparer entre elles.

Les caractéristiques de position

Nous envisagerons d'abord les caractéristiques de position, appelées ainsi, car elles localisent les valeurs des termes de la série. Les principales sont la médiane, les quartiles, les moyennes, le mode.

La médiane

La médiane est une quantité M telle qu'il existe autant de termes de la série inférieure à M que de termes supérieurs à M. Il découle de cette définition qu'une série à caractère discontinu a rarement une médiane. En effet, reprenons l'exemple de la série 1. Cette série comprend 40 termes. Ces derniers étant rangés par ordre de grandeurs croissantes, la médiane M est comprise entre la valeur du 20^e terme et celle du 21^e. Reportons-nous au tableau donnant les effectifs cumulés. Le 20^e terme est égal à 2, ainsi que le 21^e. La médiane serait donc égale à 2, mais nous constatons qu'il y a 11 termes inférieurs à 2 et 14 termes supérieurs à 2. Cette valeur de M ne correspond donc pas à la définition de la médiane.

Pour déterminer la médiane d'une série à caractère continu, on cherche le rang de la médiane en divisant l'effectif total par 2. Dans la série 2 ◆, la médiane M serait toute valeur comprise entre le 20^e et le 21^e terme. On peut confondre cette médiane avec l'un ou l'autre de ces termes. L'examen des effectifs cumulés nous apprend que le 20^e terme se trouve dans la quatrième classe. Sa valeur est donc comprise entre 35 et 40.
On peut en donner une valeur approchée (d'autant plus exacte que

◆ Voir page 459.

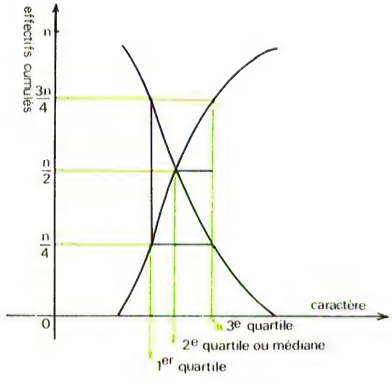
l'effectif de classe est plus grand) en supposant que tous les termes de la quatrième classe se succèdent en valeur par intervalles égaux. On cherche alors le rang de la médiane dans la classe : $20 - 13 = 7$. On calcule l'intervalle entre chaque terme de la classe en divisant l'intervalle de classe par l'effectif de classe soit 5/9. La médiane est alors égale à la limite inférieure de la classe, à laquelle on ajoute sept fois l'intervalle : 5/9, soit $35 + \frac{7 \times 5}{9} = 38{,}8$.

La médiane se détermine graphiquement en construisant les diagrammes cumulatifs. C'est l'abscisse des points de rencontre des deux courbes. En effet, ce point appartient à la courbe croissante ; son ordonnée indique le nombre de termes inférieurs à son abscisse, mais il appartient aussi à la courbe décroissante et son ordonnée indique le nombre de termes supérieurs à son abscisse. L'abscisse de ce point répond donc à la définition de la médiane ◆.

◆ Voir courbes, pages 462.

Les quartiles

Les quartiles sont des paramètres qui divisent la série en quatre parties, les termes de cette série étant ordonnés. Un quart des termes doivent être inférieurs au premier quartile ; la moitié des termes doivent être inférieurs au deuxième quartile, qui n'est autre que la médiane ; les trois quarts des termes doivent être inférieurs au troisième quartile.



Le calcul de ces quartiles est analogue à celui de la médiane. On les obtient graphiquement en cherchant sur les diagrammes cumulatifs les abscisses des points dont les ordonnées sont égales à 1/4, 1/2 et 3/4 de l'effectif total.

La moyenne arithmétique

La moyenne arithmétique est un paramètre très employé. Elle est égale à la somme des termes divisée par leur nombre.

Si l'on appelle x_i un terme quelconque de la série, la moyenne $\bar{x}$ est égale à : $\bar{x} = \frac{\sum_{1}^{n} x_i}{n}$ ◆.

◆ n est l'effectif total de la série.

$\sum_{1}^{n} x_i$ remplace la somme $x_1 + x_2 + x_3 + \ldots + x_n$.

Le calcul de la moyenne d'une série à caractère discontinu ne pose aucun problème puisque l'on connaît la valeur exacte des termes.

Reprenons la série 1.
Calcul de la moyenne :
X_i désigne les valeurs des classes ;
n_i désigne les effectifs de classe ;

$$\bar{x} = \frac{91}{40} = 2{,}02$$

X_i	n_i	$X_i n_i$
1	11	11
2	15	30
3	8	24
4	4	16
5	2	10
	40	91

Dans le cas de la série à caractère continu, on ne connaît plus la valeur exacte des termes. On suppose tous les termes d'une même classe égaux à la valeur centrale de la classe.

Nous désignons cette valeur par X_i.

$$\bar{x} = \frac{\sum_{1}^{m} X_i n_i}{n}$$

où il y a m classes.

Calcul de la moyenne de la série 2 :

$$\bar{x} = \frac{1680}{40} = 42$$

X_i	n_i	$X_i n_i$
22,5	2	45
27,5	6	165
32,5	5	162,5
37,5	9	337,5
45	7	315
55	6	330
65	5	325
	40	1 680

Signalons que l'on peut considérer ces moyennes comme des moyennes pondérées. En effet, la moyenne pondérée de plusieurs termes s'obtient en multipliant chacun d'eux par son *coefficient de pondération* ; on ajoute les produits obtenus et on divise le résultat par la somme des coefficients de pondération.
Soient m termes : $X_1, X_2, \ldots, X_m$ affectés des coefficients de pondération $n_1, n_2, \ldots, n_m$.
La moyenne pondérée est :

$$\bar{x} = \frac{X_1 n_1 + X_2 n_2 + \ldots + X_m n_m}{n_1 + n_2 + \ldots + n_m} \qquad \bar{x} = \frac{\sum_{1}^{m} X_i n_i}{\sum_{1}^{m} n_i}$$

En réalité, une moyenne arithmétique pondérée est une moyenne arithmétique ordinaire si l'on suppose que chaque terme est répété autant de fois qu'il y a d'unités dans son coefficient de pondération.
Naturellement, le coefficient de pondération doit être un nombre entier, ce qui, d'ailleurs, est toujours réalisable.
La médiane et la moyenne ont l'avantage de caractériser par un

seul nombre l'ensemble des termes d'une série. On utilisera de préférence la médiane dans le cas où la plupart des valeurs des termes sont assez voisines, quelques-unes seulement s'en éloignant beaucoup.
Par exemple : dans une entreprise, il y a 1 000 ouvriers dont les salaires sont assez voisins et une vingtaine de techniciens qui ont un salaire nettement supérieur. On voit que ces derniers n'influenceront pas la médiane des salaires, alors qu'ils modifieront beaucoup la moyenne.

Il existe des moyennes autres que la moyenne arithmétique, mais employées beaucoup plus rarement.

La moyenne harmonique de plusieurs nombres est l'inverse de la moyenne arithmétique des inverses de ces nombres◆.
Soit x_h la moyenne harmonique des nombres $x_1, x_2, \ldots, x_n$:

◆ On rencontre l'utilisation de cette moyenne dans les calculs d'indice.

$$x_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}$$

La moyenne géométrique◆ des nombres précédents serait :

◆ On utilise cette moyenne pour caractériser une série de rapports représentant les variations relatives d'un phénomène.

$$x_g = \sqrt[n]{x_1 \times x_2 \times \ldots \times x_n}$$

Exemple : nombre d'objets fabriqués par différentes usines au cours de deux années consécutives, en milliers :

1re année	100	80	120	90	75
2e année	120	125	150	100	80

Moyenne géométrique des rapports :

$$x_g = \sqrt[5]{\frac{120}{100} \times \frac{125}{80} \times \frac{150}{120} \times \frac{100}{90} \times \frac{80}{75}} = 1{,}226$$

La moyenne géométrique est réversible, car, si, l'on calculait la moyenne géométrique des inverses des rapports précédents, on trouverait : 1/1,226.
La moyenne arithmétique n'a pas cette propriété.

Le mode

Un dernier paramètre de position est le mode.

Dans une série à caractère discontinu, le mode est la valeur du caractère qui correspond à la classe renfermant le plus grand nombre de termes. Dans la série 1, le mode est égal à 2◆.

◆ Voir page 458.

Dans une série à caractère continu, le mode est égal à la valeur centrale de la classe à laquelle correspond la plus grande densité.
Dans la série 2, le mode est égal à 37,5.
Le mode renseigne sur la valeur du caractère la plus fréquente.

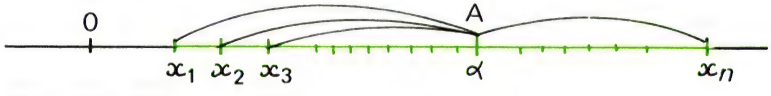
Les caractéristiques de dispersion

Deux séries statistiques peuvent avoir la même moyenne ou la même médiane et ne pas se ressembler du tout, si elles présentent des « dispersions » différentes autour de ces valeurs ; par exemple, deux séries de salaires dont la moyenne mensuelle est de 1 200 F, les salaires de la première série variant entre 1 100 et 1 400 F et ceux de la deuxième variant entre 700 et 5 000 F. Il est donc indispensable, pour caractériser une série statistique, d'ajouter à la notion de paramètre de position celle de paramètre de dispersion.

Écart

Nous essaierons de définir d'abord le mot « écart », qui est si souvent employé en statistique.
Soit α un nombre arbitraire : la différence entre un terme de la série et le nombre α s'appelle écart à α. Chaque terme de la série peut être considéré comme l'écart au nombre 0. Exprimer les termes de la série par leurs écarts à α revient à faire un changement d'origine.

Termes de la série disposés sur un axe d'origine O.



$\overline{Ax_1}$, $\overline{Ax_2}$, $\overline{Ax_3}$, ..., $\overline{Ax_n}$: écarts au nombre α.

On est ainsi amené à considérer : la série des écarts à la médiane, la série des écarts à la moyenne.

Propriétés des écarts

On démontre très facilement : la somme des écarts à la médiane est plus petite que la somme des écarts à n'importe quel autre nombre ; la somme des écarts à la moyenne est nulle ; la somme des carrés des écarts à la moyenne est plus petite que la somme des carrés des écarts à tout autre nombre.
La dispersion d'une série est caractérisée par la grandeur des écarts à une valeur centrale (médiane ou moyenne). Plus l'écart est grand, plus la série s'étale autour de cette valeur centrale. La plus utilisée de ces valeurs est la moyenne. La somme des écarts à la moyenne

étant nulle, on calcule la somme de la valeur absolue des écarts ou encore la somme du carré des écarts.

Écart médian, variance, écart type

La moyenne de la valeur absolue des écarts se nomme écart médian ; la moyenne du carré des écarts se nomme variance ou fluctuation. Sa racine carrée est l'écart quadratique moyen ou écart type noté σ :

$$\sigma^2 = \frac{\sum_1^n (x_i - \bar{x})^2}{n} \blacklozenge$$

◆ $\sum_1^n (x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2$,

ou encore, en utilisant les valeurs centrales des classes X_i :

$$\sigma^2 = \frac{\sum_1^m (X_i - \bar{x})^2 n_i}{n}$$

Si σ est petit, cela signifie que les termes de la série sont peu dispersés autour de la moyenne ou, tout au moins, que les termes très éloignés de cette moyenne sont peu nombreux.
L'écart type étant d'un emploi fondamental en statistique, nous en donnerons un exemple de calcul.
Reprenons la série 2 dont la moyenne est 42.

X_i	n_i	$X_i - \bar{x}$	$(X_i - \bar{x})^2$	$(X_i - \bar{x})^2 n_i$
22,5	2	− 19,5	380,25	760,5
27,5	6	− 14,5	210,25	1 261,5
32,5	5	− 9,5	90,25	451,25
37,5	9	− 4,5	20,25	182,25
45	7	3	9	63
55	6	13	169	1 014
65	5	23	529	2 645
	40			6 377,50

$$\sigma^2 = \frac{6\,377{,}5}{40} = 159{,}37 \qquad \sigma = \sqrt{159{,}37} = 12{,}62$$

Afin de montrer l'importance de l'écart type pour caractériser la dispersion d'une série ◆, nous allons étudier un exemple particulier, où la série statistique se rapproche d'une courbe normale ou courbe de Gauss.

◆ Dans cette série, les valeurs des termes s'échelonnent en gros entre 22 5 et 65. On constate que ces nombres correspondent environ à $42 \pm 2\,\sigma$ (42 étant la valeur moyenne), soit :
$42 \pm 2 \times 12{,}62$
$42 \pm 25{,}24$
17 et 67.

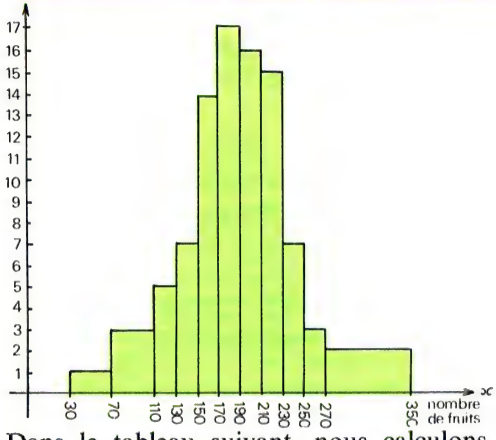
Courbe de Gauss

Soit la série statistique « répartition de 10 000 arbres fruitiers d'après le nombre de fruits produits en une saison ». Le nombre des arbres est exprimé en pourcentage pour simplifier les calculs.

fruits	30 à <70	70 à <110	110 à <130	130 à <150	150 à <170	170 à <190	190 à <210	210 à <230	230 à <250	250 à <270	270 à <350	total
arbres	2	6	5	7	14	17	16	15	7	3	8	100

La représentation graphique de cette série est un histogramme. Pour le tracer, nous calculerons les hauteurs des rectangles (voir le calcul page 461 : 20 étant le plus petit intervalle de classe, c'est lui qui sert d'unité).

intervalle de classe	40_2	40_2	20_1	20_1	20_1	20_1	20_1	20_1	20_1	20_1	80_4
hauteur des rectangles	1	3	5	7	14	17	16	15	7	3	2



Dans le tableau suivant, nous calculons la moyenne, la variance et l'écart type de la série.

val. centrales des classes X_i	n_i	$X_i n_i$	$X_i - \bar{x}$	$(X_i - \bar{x})^2$	$(X_i - \bar{x})^2 n_i$
50	2	100	− 139,6	19 488,16	38 976,32
90	6	540	− 99,6	9 920,16	59 520,96
120	5	600	− 69,6	4 844,16	24 220,80
140	7	980	− 49,6	2 460,16	17 221,12
160	14	2 240	− 29,6	876,16	12 266,24
180	17	3 060	− 9,6	92,16	1 566,72
200	16	3 200	10,4	108,16	1 730,56
220	15	3 300	30,4	924,16	13 862,40
240	7	1 680	50,4	2 540,16	17 781,12
260	3	780	70,4	4 956,16	14 868,48
310	8	2 480	120,4	14 496,16	115 969,28
		18 960			317 984

La moyenne

$$\bar{x} = \frac{\Sigma\, x_i\, n_i}{n} = \frac{18\,960}{100} = 189{,}60$$

$$\sigma^2 = \frac{317\,984}{100} = 3\,179{,}84$$

$$\sigma = 56{,}39$$

Il est intéressant de mesurer la dispersion des observations autour de la moyenne à l'aide de l'écart type.

Par exemple : combien y a-t-il d'arbres fruitiers dont la production est comprise entre $\bar{x} + \sigma$ et $\bar{x} - \sigma$, c'est-à-dire entre :

$$189{,}6 + 56{,}39 = 246 \text{ et } 189{,}6 - 56{,}39 = 133{,}2 ?$$

Il faut additionner les effectifs de classe dont les valeurs sont comprises entre 133 et 246, soit environ :

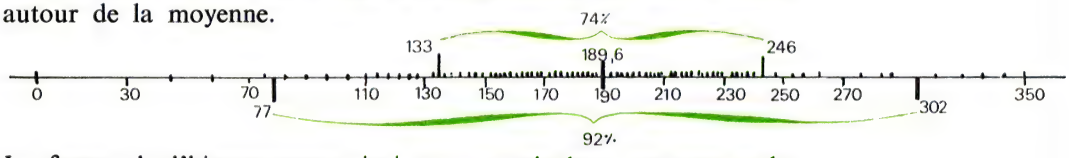
$$6 \blacklozenge + 14 + 17 + 16 + 15 + 6 \blacklozenge = 74$$

◆ Ces deux chiffres sont les effectifs de la 4e et 9e classe, diminués légèrement pour les ajuster à 133 et à 246 (la 4e étant 130 et la 9e 250).

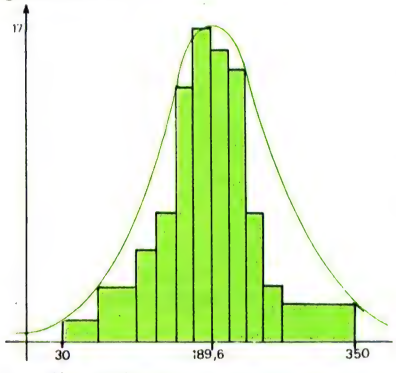
74 arbres sur 100 produisent entre 133 et 246 fruits.
Combien y a-t-il d'arbres fruitiers dont la production est comprise entre $\bar{x} + 2\sigma$ et $\bar{x} - 2\sigma$, soit entre $189{,}6 + 112{,}8 \simeq 302$ et $189{,}6 - 112{,}8 \simeq 77$?
De la même façon, nous additionnons les effectifs de classe. Nous en prenons 5 dans la deuxième classe et 3 dans la dernière classe.

$$5 + 5 + 7 + 14 + 17 + 16 + 15 + 7 + 3 + 3 = 92$$

92 arbres sur 100 produisent entre 70 et 300 fruits.
En représentant chaque élément de la série par des points sur un axe, on peut donner une image de la répartition des observations autour de la moyenne.



La forme de l'histogramme ainsi que ces résultats concernant la dispersion indiquent que nous avons affaire à une répartition gaussienne.



La figure ci-contre montre que l'histogramme pourrait s'ajuster à une courbe de Gauss dont il présente les principaux caractères : forme, symétrie, concentration des observations autour de la moyenne.

La dispersion des observations autour de la moyenne est aussi une dispersion de Gauss, ainsi que nous pouvons le vérifier en consultant une table de Gauss.
Pour cela, il faut d'abord calculer les valeurs de la variable réduite ◆.

◆ Ces notions sont reprises plus loin, page 496.

Appelons x le nombre de fruits produits par les arbres. Cette variable x prend les valeurs indiquées par le tableau de la série, c'est-à-dire que, dans la première classe, x prend des valeurs comprises entre 30 et 70, dans la deuxième classe, des valeurs comprises entre 70 et 110, etc.
A chaque valeur de x correspond une valeur de la variable réduite qui est égale à : $\frac{x - \bar{x}}{\sigma}$, soit $t = \frac{x - 189{,}6}{56{,}39}$.

Remarquons que remplacer la variable x par la variable réduite t revient à faire un changement d'origine (la valeur moyenne de x étant prise pour origine) et un changement d'unité (on considère σ comme nouvelle unité pour mesurer la production des arbres, soit un panier de 56,39 fruits).
Calcul de la variable réduite correspondant aux limites de classe : voir tableau ci-contre.

Si $x = 189{,}6$ $\quad t = 0$.
Si $x = 189{,}6 + 56{,}39$ $\quad t = 1$.
si $x = 189{,}6 - 56{,}39$ $\quad t = -1$.

Lorsque x est compris entre 189,6 — 56,39 et 189,6 + 56,39, t est compris entre — 1 et + 1.
Lorsque x est compris entre 189,6 — 2 × 56,39 et 189,6 + 2 × 56,39, t est compris entre — 2 et + 2.
La table de Gauss donne les probabilités pour que les valeurs de la variable réduite soit inférieures à une valeur donnée◆.
Les renseignements suivants se déduisent des indications de la table pour $t = 1$ et $t = 2$: dans une répartition gaussienne, 68 % des observations sont comprises entre $\bar{x} - \sigma$ et $\bar{x} + \sigma$; 95 % des observations sont comprises entre $\bar{x} - 2\sigma$ et $\bar{x} + 2\sigma$, ce qui signifie que 68 arbres sur 100 doivent fournir entre 133 et 246 fruits et que 95 arbres sur 100 doivent fournir entre 77 et 302 fruits pour que la répartition soit gaussienne. Ces résultats sont voisins de ceux donnés par la série.

$x = 30$	$t = -2{,}83$
$x = 70$	$t = -2{,}12$
$x = 110$	$t = -1{,}41$
$x = 150$	$t = -0{,}70$
$x = 170$	$t = -0{,}35$
$x = 190$	$t = 0{,}007$
$x = 210$	$t = 0{,}36$
$x = 230$	$t = 0{,}72$
$x = 250$	$t = 1{,}07$
$x = 270$	$t = 1{,}42$
$x = 350$	$t = 2{,}84$

◆ Nous donnons un extrait de cette table (page 499) ainsi que la façon de l'utiliser.

On peut penser que cette production est « normale ».
En effet, les phénomènes qui se distribuent selon une courbe de Gauss ou « courbe normale » sont ceux qui sont dus au hasard ou, tout au moins, qui sont le résultat d'un grand nombre de facteurs agissant tous indépendamment les uns des autres.

SÉRIE STATISTIQUE DOUBLE

Il peut arriver que l'on étudie une population sous deux caractères. *Par exemple, l'ensemble des propriétés rurales déjà envisagé, étudié simultanément sous les deux caractères « nombre de personnes actives » et « surface cultivable ».*

L'enquête consiste alors à mesurer sur chaque élément de la population les deux caractères. On obtient un ensemble de couples, appelé série statistique double :

(1,20)	(2,42)	(1,25)	(3,48)	(2,38)	(1,35)	(2,50)	(2,35)
(2,35)	(2,40)	(1,26)	(3,52)	(3,48)	(1,30)	(2,27)	(2,36)
(1,22)	(3,39)	(1,30)	(5,70)	(5,68)	(2,36)	(2,32)	(2,42)
(3,45)	(4,65)	(2,35)	(1,25)	(1,30)	(2,38)	(3,55)	(3,52)
(4,60)	(4,70)	(4,58)	(1,26)	(1,26)	(2,45)	(3,56)	(2,30)

Le premier élément du couple représente le nombre de personnes actives dans la propriété, le deuxième donnant la surface cultivable. Pour ordonner cette série, on range les couples par ordre de valeurs croissantes du premier élément, sans tenir compte du deuxième :

(1,20)	(1,25)	(1,30)	(2,38)	(2,27)	(2,30)	(3,48)	(4,60)
(1,22)	(1,26)	(2,35)	(2,36)	(2,32)	(3,45)	(3,55)	(4,65)
(1,25)	(1,30)	(2,42)	(2,38)	(2,35)	(3,39)	(3,56)	(4,70)
(1,26)	(1,26)	(2,40)	(2,45)	(2,36)	(3,48)	(3,52)	(5,70)
(1,30)	(1,35)	(2,35)	(2,50)	(2,42)	(3,52)	(4,58)	(5,68)

On partage l'ensemble précédent en classes, en ne considérant encore que le premier caractère. On obtient ainsi le tableau :

personnes actives	1	2	3	4	5
propriétés	11	15	8	4	2

A l'intérieur de chaque classe, on ordonne alors les éléments d'après le second caractère :

1	2	3	4	5
1,20	2,37	3,39	4,58	5,68
1,22	2,30	3,45	4,60	5,70
1,25	2,32	3,48	4,65	
1,25	2,35	3,48	4,70	
1,26	2,35	3,52		
1,26	2,35	3,52		
1,26	2,36	3,55		
1,30	2,36	3,56		
1,30	2,37			
1,30	2,38			
1,35	2,40			
	2,42			
	2,42			
	2,45			
	2,50			

Tableau de corrélation

Il reste à former des classes en considérant le second caractère, et l'on obtient ainsi un tableau à double entrée, appelé tableau de corrélation :

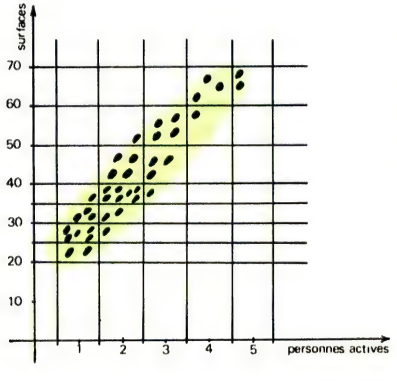
surfaces \ personnes actives	1	2	3	4	5	total
20 à < 25	2					2
25 à < 30	5	1				6
30 à < 35	3	2				5
35 à < 40	1	7	1			9
40 à < 50		4	3			7
50 à < 60		1	4	1		6
60 à < 70				3	2	5
total	11	15	8	4	2	40

Chaque colonne et chaque ligne du tableau constituent une série simple, c'est-à-dire une répartition de la population sous un seul caractère.

La dernière colonne et la dernière ligne donnent la répartition de la population entière sous les caractères : surface cultivable (colonne), nombre de personnes actives (ligne). On les appelle distributions marginales. Dans toutes les autres colonnes et toutes les autres lignes se trouvent des distributions conditionnelles, c'est-à-dire la répartition d'une partie de la population sous un caractère, cette partie étant conditionnée par l'autre caractère. Par exemple, ligne 3 : répartition des propriétés dont la surface est comprise entre 30 et 35 ha selon le nombre de personnes actives ; il y a trois propriétés qui ont chacune une personne active, et deux qui ont chacune deux personnes actives.

Nuage de points

La série double est représentée graphiquement par un nuage de points.



Pour placer ces points, on forme un quadrillage. On porte en abscisse les valeurs du caractère « nombre de personnes actives » et en ordonnée celles du caractère « surface cultivable ». Puis on trace des parallèles aux axes ainsi que l'indique la figure.

On forme ainsi des cases dans lesquelles on répartit « au hasard » un nombre de points égal au nombre d'éléments de la population indiqué par le tableau de corrélation.

Corrélation

La forme du nuage donne un renseignement précieux sur la corrélation entre les deux caractères.

On dit qu'il y a corrélation entre deux caractères des éléments d'un même ensemble lorsque les variations de l'un sont dépendantes des variations de l'autre. Cette dépendance est intermédiaire entre la liaison fonctionnelle et l'indépendance complète.

La liaison fonctionnelle est exprimée par une relation mathématique qui permet de déterminer exactement la valeur d'un caractère en fonction des valeurs de l'autre.

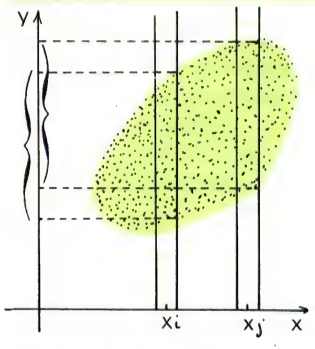
Exemple : soit une population de ballons envisagée sous les deux caractères « rayon » et « volume ». On mesure sur chaque ballon la valeur de ces deux caractères. Ils sont liés entre eux par une relation fonctionnelle :

$$V = \frac{4}{3}\pi R^3 \quad (\text{où } V = \text{volume et } R = \text{rayon})$$

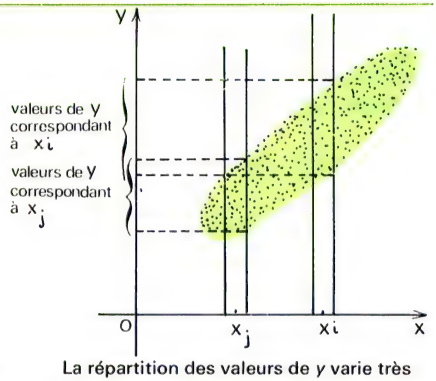
L'indépendance est réalisée lorsque les variations de valeurs de l'un des caractères ne sont en aucun cas la conséquence des valeurs de l'autre.

Exemple : dans la population de ballons envisagée plus haut, le volume du ballon et sa couleur sont deux caractères indépendants.

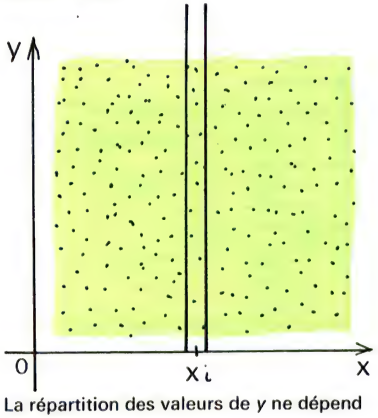
Si maintenant, dans une population d'adultes masculins du même âge, on mesure la taille et le poids, on constate généralement que plus la taille augmente, plus le poids augmente et que, de plus, dans un groupe d'individus de même taille, la variance de la série des poids est moindre que la variance de la série des poids de la population totale.
Il existe donc une certaine dépendance entre les deux caractères : une liaison statistique. On dit que les deux caractères sont en corrélation.
Cette liaison est mise en évidence par la forme du nuage de points qui représente graphiquement la série statistique double.
Lorsque le nuage de points affecte la forme d'une ellipse, le grand axe de cette ellipse étant dirigé obliquement par rapport aux axes, les deux caractères sont en corrélation linéaire, corrélation d'autant plus serrée que l'ellipse est plus aplatie.



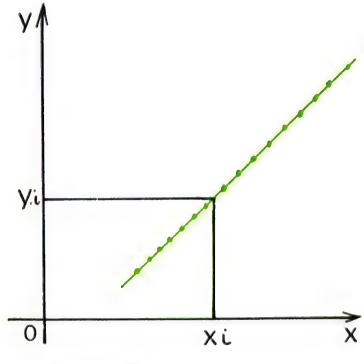
La répartition des valeurs de *y* varie peu en même temps que *x* varie : **liaison large.**



La répartition des valeurs de *y* varie très nettement en même temps que celles de *x* : **liaison serrée.**

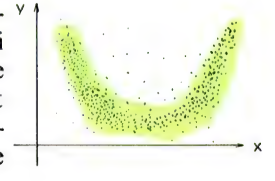


La répartition des valeurs de *y* ne dépend pas des valeurs de *x* : **indépendance totale.**



y est fonction de *x* : **liaison fonctionnelle.**

Il peut arriver que le nuage de points ait une densité forte au voisinage d'une courbe et très faible partout ailleurs. La liaison qui existe entre les deux caractères n'est plus linéaire, mais il semble que l'on pourrait ajuster ces points à une courbe dont on calculerait l'équation algébrique et qui nous permettrait de déterminer approximativement la valeur d'un caractère, connaissant les valeurs de l'autre.



A partir de ces petits schémas, on s'aperçoit que la corrélation pose un double problème :

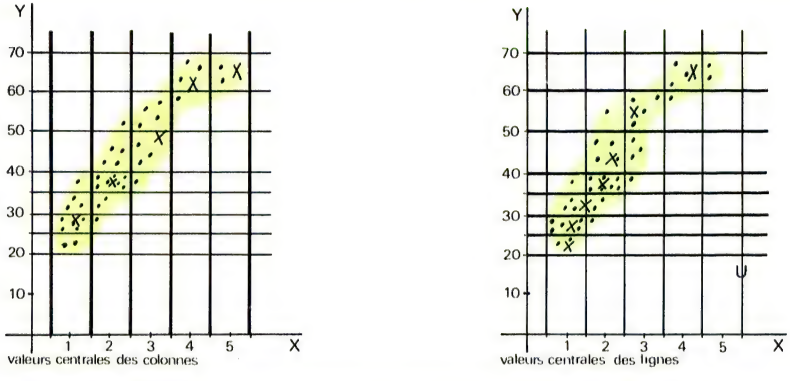
— trouver un indice mesurant le degré de liaison entre les deux caractères ;

— préciser la forme de cette liaison (linéaire ou autre) ; si possible, établir une relation mathématique entre les deux caractères, per-

mettant de trouver les valeurs de l'un d'entre eux si les valeurs de l'autre sont connues.
En réalité, ces deux problèmes sont liés et la formule donnant l'indice de corrélation varie avec la forme mathématique de la relation entre les deux caractères.

Corrélation linéaire

Reprenons l'exemple des fermes étudiées sous les deux caractères « surfaces cultivables » et « nombre de personnes actives ». Marquons d'une croix les valeurs centrales des bandes parallèles à l'axe des ordonnées et les valeurs centrales des bandes parallèles à l'axe des abscisses :



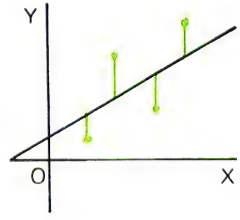
Dans les deux cas, les croix adoptent une direction linéaire. La première série de croix donne par leurs coordonnées une valeur approximative de Y en fonction de X.
Si nous ajustons cette série de croix à une droite, l'équation de la droite $Y = f(X)$ donnera la surface cultivable des fermes en fonction des nombres de personnes actives. C'est la *droite de régression de* Y *en* X.

De même, la deuxième série de croix donne par leurs coordonnées la valeur approximative de X en fonction de Y. Si nous ajustons la deuxième série de croix à une droite, l'équation de cette droite $X = f(Y)$ donnera le nombre de personnes actives à la ferme en fonction de la surface cultivable. C'est la *droite de régression de* X *en* Y.
Nous allons établir l'équation de ces deux droites, dans le cas de notre exemple.

Méthode des moindres carrés

L'ajustement des points à une droite se fait par la méthode des moindres carrés. Cette méthode consiste à trouver l'équation d'une droite telle que la somme des carrés des distances des points à la droite soit minimale.

Si l'on compte les distances parallèlement à OY, l'équation de la droite d'ajustement est la suivante :



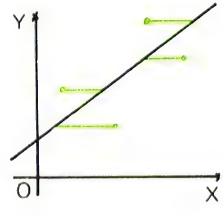
$$Y - \overline{Y} = k(X - \overline{X})$$

où $\overline{Y}$ est la moyenne des ordonnées des points (moyenne de la colonne marginale) et $\overline{X}$ la moyenne des abscisses des points (moyenne de la ligne marginale).

$$k = \frac{\Sigma (X_i - \overline{X})(Y_i - \overline{Y})}{\Sigma (X_i - \overline{X})^2}$$

Cette équation sera celle de la droite de régression de Y en X.

Si l'on mesure la distance des points à la droite parallèlement à l'axe des abscisses, l'équation de la droite d'ajustement est la suivante :



$$X - \overline{X} = k'(Y - \overline{Y})$$

$$k' = \frac{\Sigma (X_i - \overline{X})(Y_i - \overline{Y})}{\Sigma (Y_i - \overline{Y})^2}$$

C'est l'équation de la droite de régression de X en Y.

Pour établir les équations de ces droites, nous calculerons successivement :

$\overline{X}$: moyenne des abscisses des points ;

$\overline{Y}$: moyenne des ordonnées des points ;

$X_i - \overline{X}$, $Y_i - \overline{Y}$: nouvelles coordonnées des points en prenant comme origine le point de coordonnées $\overline{X}$, $\overline{Y}$;

$\Sigma (X_i - \overline{X})(Y_i - \overline{Y})$: somme des produits des nouvelles coordonnées de chaque point ;

$\Sigma (X_i - \overline{X})^2$: somme des carrés des nouvelles abscisses ;

$\Sigma (Y_i - \overline{Y})^2$: somme des carrés des nouvelles ordonnées.

Les calculs sont souvent présentés dans un tableau semblable à celui de la page suivante.

1	2		3	4	5	6	7	8	9	10	11	12	13	14
personnes actives			1	2	3	4	5	totaux n_i	$Y_i n_i$	$Y_i - \bar{Y}$	$(X_i - \bar{X})\, a_i$	$(X_i - \bar{X})\, a_i\, (Y_i - Y)$	$(Y_i - \bar{Y})^2$	$(Y_i - \bar{Y})^2\, n_i$
	surfaces	Classes valeurs centrales												
1	20 à 25	22,5	2					2	45	-19,5	- 2,55	49,725	380,25	760,5
2	25 à 30	27,5	5	1				6	165	-14,5	- 6,65	96,425	210,25	1 261,5
3	30 à 35	32,5	3	2				5	162,5	- 9,5	- 4,375	41,562 5	90,25	451,25
4	35 à 40	37,5	1	7	1			9	337,5	- 4,5	- 2,475	11,137 5	20,25	182,25
5	40 à 50	45		4	3			7	315	3	1,175	3,525	9	63
6	50 à 60	55		1	4	1		6	330	13	4,35	56,55	169	1 014
7	60 à 70	65				3	2	5	325	23	10,625	244,375	529	2 645
8	totaux n'_i		11	15	8	4	2	40	1 680			503		6 377,5
9	$X_i \times n'_i$		11	30	24	16	10	91						
10	$X_i - \bar{X}$		-1,275	-0,275	0,725	1,725	2,725							
11	$(Y_i - \bar{Y})\, a_i$		-144,5	-40	56,5	82	46							
12	$(Y_i - \bar{Y})\, a_i\, (X_i - X)$		184,237 5	11	40,962 5	141,45	125,35	503						
13	$(X_i - \bar{X})^2$		1,625	0,075	0,525	2,975	7,425							
14	$(X_i - \bar{X})^2\, n'_i$		17,875	1,125	4,2	11,9	14,85	49,95						

$\bar{x} = 2,275$
$\bar{y} = 42$

Explication des calculs (suivre sur le graphique de la page 476).

1) Calcul de $\overline{X}$:

il y a 11 points d'abscisse 1 (1re colonne du graphique) ;
— 15 — 2 (2e —) ;
— 8 — 3 (3e —) ;
— 4 — 4 (4e —) ;
— 2 — 5 (5e —).

Dans la 9e ligne du tableau, on a fait les produits :

11×1; 15×2; 8×3; 4×4; 2×5

Le total de ces produits est enregistré au croisement de la colonne 8 et de la ligne 9 :

$$\overline{X} = \frac{\Sigma\, X_i\, n_i}{\Sigma\, n_i} = \frac{91}{40} = 2{,}275$$

2) Calcul de $\overline{Y}$:

il y a 2 points d'ordonnée 22,5 (1re ligne du graphique) ;
— 6 — 27,5 (2e —) ;
— 5 — 32,5 (3e —) ;
— 9 — 37,5 (4e —) ;
— 7 — 45 (5e —) ;
— 6 — 55 (6e —) ;
— 5 — 65 (7e —).

Dans la 9e colonne du tableau, on a fait les produits :

$22{,}5 \times 2$; $27{,}5 \times 6$; $32{,}5 \times 5$; $37{,}5 \times 9$; 45×7; 55×6; 65×5

Le total de ces produits est enregistré au croisement de la colonne 9 et de la ligne 8

$$\overline{Y} = \frac{\Sigma\, Y_i\, n_i}{\Sigma\, n_i} = \frac{1\,680}{40} = 42$$

3) Calcul des écarts :

$X_i - \overline{X}$: ligne 10 du tableau ;

$Y_i - \overline{Y}$: colonne 10 du tableau.

4) Calcul des produits $(X_i - \overline{X})\,(Y_i - \overline{Y})$

Considérons les points de la colonne 1 (sur le graphique). La somme des produits $(X_i - \overline{X})\,(Y_i - \overline{Y})$ relatifs à ces points est égale à :

2 fois $(-1{,}275) \times (-19{,}5) + 5$ fois $(-1{,}275) \times (-14{,}5)$
$+ 3$ fois $(-1{,}275) \times (-9{,}5) + 1$ fois $(-1{,}275) \times (-4{,}5)$

Soit, en mettant — 1,275 en facteur :

$$(-1{,}275)\,[2 \times (-19{,}5) + 5 \times (-14{,}5) + 3 \times (-9{,}5) + (-4{,}5)] = (-1{,}275) \times (-144{,}5)$$

Le résultat de la quantité entre crochets est inscrit dans la première case de la ligne 11 du tableau.
De même pour les points de la colonne 2 (du graphique) :

$$1 \text{ fois } (-0{,}275) \times (-14{,}5) + 2 \text{ fois } (-0{,}275) \times (-9{,}5) + 7 \text{ fois } (-0{,}275) \times (-4{,}5) + 4 \text{ fois } (-0{,}275) \times 3 + 1 \text{ fois } (-0{,}275) \times 13$$

Soit :

$$(-0{,}275)\,[-14{,}5 + 2 \times (-9{,}5) + 7\,(-4{,}5) + 4 \times 3 + 13] = (-0{,}275) \times (-40)$$

Le résultat de la quantité entre crochets est inscrit dans la 2e case de la ligne 11.
On procède de la même manière pour les points de chacune des colonnes.
Il suffit alors de multiplier les résultats inscrits dans la ligne 11 par les valeurs correspondantes de $(X_i - \overline{X})$ inscrites dans la ligne 10.
$\Sigma\,(X_i - \overline{X})\,(Y_i - \overline{Y})$ est enregistré au croisement des ligne 12 et colonne 8.

On reprend souvent le calcul à titre de vérification et en faisant une autre mise en facteur. On considère cette fois les points d'une même ligne (sur le graphique) et l'on met en facteur la valeur de Y_i correspondante. Ces calculs sont faits colonnes 11 et 12.
Calcul de $(X_i - \overline{X})^2$: ligne 13.
On multiplie chaque carré par le nombre de points de chaque colonne (du graphique) : ligne 14 du tableau.
Le total est inscrit au croisement de la ligne 14 et de la colonne 8.
Calcul de $(Y_i - \overline{Y})^2$: même chose, voir colonnes 13 et 14.

Équation des droites de régression

$$Y - \overline{Y} = k\,(X_i - \overline{X})$$

$$k = \frac{503}{49{,}95} = 10{,}07$$

$$Y - 42 = 10{,}07\,(X - 1{,}275) \qquad Y = 10{,}07\,X + 29{,}16$$

$$X - \overline{X} = k'\,(Y_i - \overline{Y})$$

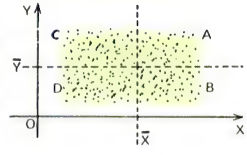
$$k' = \frac{503}{6\,377{,}5} = 0{,}078$$

$$X - 1{,}275 = 0{,}078\,(Y - 42) \qquad X = 0{,}078\,Y - 2$$

Coefficient de corrélation

Le degré de corrélation entre deux caractères d'une même population sera caractérisé par l'écartement des droites. En effet, si les deux caractères étaient liés par une liaison fonctionnelle, les deux droites seraient confondues.

Si les deux caractères étaient complètement indépendants, les deux droites seraient respectivement parallèles aux axes, car la valeur de $\Sigma (X_i - \overline{X})(Y_i - \overline{Y})$ serait sensiblement nulle, la somme des produits relatifs aux points de la région A et B annulant respectivement la somme des produits relatifs aux points de la région C et D.



Dans ce cas, les équations des droites de régression seraient $X = \overline{X}$ et $Y = \overline{Y}$.

Nous voyons que, lorsque nous passons de la relation fonctionnelle à l'indépendance totale, les deux droites, d'abord confondues, tendent à devenir perpendiculaires. Il semble donc que le degré de liaison peut être mesuré par l'écartement plus ou moins grand des droites, cet écartement étant par ailleurs mesuré numériquement par le produit des pentes des deux droites de corrélation. On appelle coefficient de corrélation linéaire r la racine carrée de ce produit :

$$r^2 = k \times k'$$

$$r^2 = \frac{\Sigma (X_i - \overline{X})(Y_i - \overline{Y})}{\Sigma (X_i - X)^2} \times \frac{\Sigma (X_i - \overline{X})(Y_i - \overline{Y})}{\Sigma (Y_i - \overline{Y})^2}$$

Dans notre exemple, $r^2 = 10{,}07 \times 0{,}078 = 0{,}78$ et $r = 0{,}88$.

Cas de la liaison fonctionnelle : les deux droites sont confondues; elles ont donc même équation, qui sont :

$$Y - \overline{Y} = k\ (X - \overline{X})$$

et $X - \overline{X} = k'(Y - \overline{Y})$ ou $Y - \overline{Y} = \frac{1}{k'}(X - \overline{X})$

donc $k = \frac{1}{k'}$ $k \times k' = 1$

et $r^2 = 1$

Cas de l'indépendance totale : les droites sont parallèles aux axes et, par suite :

$$\Sigma (X_i - \overline{X})(Y_i - \overline{Y}) = 0$$

$$r = 0$$

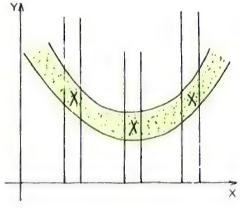
La corrélation est d'autant plus grande que k^2 se rapproche de 1 et d'autant plus faible que k^2 se rapproche de 0.

En dessous de $r = 0{,}25$, on a tendance à penser qu'il n'y a aucune corrélation entre les deux caractères.

Signe du coefficient de corrélation

On convient de donner à r le signe de $\Sigma\,(X_i - \overline{X})\,(Y_i - \overline{Y})$. Une corrélation positive correspond donc à des droites de régression croissante. Elles indiquent que X et Y varient dans le même sens. Il y a un lien de proportionnalité entre les deux caractères.
Une corrélation négative correspond à des droites de régression décroissante. Elles indiquent que X et Y varient en sens contraire. Les valeurs des deux caractères sont inversement proportionnelles.

Emploi du coefficient de corrélation



Le coefficient de corrélation ne peut être utilisé que dans le cas d'une corrélation linéaire, c'est-à-dire lorsque les valeurs centrales des bandes sont à peu près alignées.
Dans un cas comme celui de la figure ci-contre, ce serait impossible. Remarquons également que le coefficient de corrélation traduit uniquement l'existence d'une certaine liaison entre les deux caractères, l'existence de variations concomitantes, mais ne permet en aucun cas de dire si l'une des variables agit comme cause à l'égard de l'autre ou si elles dépendent toutes les deux d'une même troisième.

Rapport de corrélation de Pearson

Ce rapport peut être employé indépendamment de tout ajustement analytique.
Si l'on ne considère que la variable Y, on sait qu'on peut caractériser la dispersion de Y par sa variance :

$$\sigma_y^2 = \frac{\Sigma\,(Y - \overline{Y})^2}{n} \text{ (variance de la colonne marginale)}$$

Admettre l'hypothèse d'une liaison entre X et Y, c'est admettre que, pour chaque valeur particulière de X, c'est-à-dire dans chaque colonne du tableau de corrélation caractérisée par $X = X_i$, la dispersion se fait non plus autour de la moyenne générale $\overline{Y}$, mais autour de la moyenne des Y de cette colonne: $\overline{Y}_{X_i}$ et que la variance d'une série conditionnelle est inférieure à la variance de la population totale.
Pearson a défini un rapport de corrélation de Y en X qui tient compte de cette réduction de variance :

$$\rho^2_{Y/X} = \frac{\Sigma\, n'_i\,(\overline{Y}_{X_i} - \overline{Y})}{\Sigma\,(Y - \overline{Y})^2} \qquad 0 < \rho < 1$$

On définirait de même un rapport de corrélation de X en Y, en général différent du précédent.

On a $r = \rho$ dans le cas où la ligne des moyennes des bandes est effectivement une ligne droite. La comparaison de r et ρ peut donc donner lieu à un test de linéarité.
Le rapport de Pearson peut être utilisé dans la comparaison de deux variables quelle que soit la forme mathématique de leur liaison, l'une quantitative, l'autre qualitative. Comme il convient naturellement dans le cas de la liaison linéaire, nous allons donner un exemple de calcul sur la série double déjà étudiée.

Y \ X	1	2	3	4	5	n_i
22,5	2					2
27,5	5	1				6
32,5	3	2				5
37,5	1	7	1			9
45		4	3			7
55		1	4	1		6
65				3	2	5
n'_i	11	15	8	4	2	40
	$Y_{X=1}$	$Y_{X=2}$	$Y_{X=3}$	$Y_{X=4}$	$Y_{X=5}$	Y

Calcul de $Y_{X=1}$ (colonne 1) :

$$\overline{Y}_{X=1} = \frac{22,5 \times 2 + 27,5 \times 5 + 32,5 \times 3 + 37,5}{11} = 28,86$$

$$\overline{Y}_{X=2} = \frac{27,5 + 32,5 \times 2 + 37,5 \times 7 + 45 \times 4 + 55}{15} = 39,33$$

$$\overline{Y}_{X=3} = \frac{37,5 + 45 \times 3 + 55 \times 4}{8} = 49,06$$

$$\overline{Y}_{X=4} = \frac{55 + 65 \times 3}{4} = 62,5$$

$$\overline{Y}_{X=5} = \frac{65 \times 2}{2} = 65$$

$\overline{Y} = 42$: déjà calculé page 479.

$\Sigma (Y - \overline{Y})^2 = 6\,377,5$: déjà calculé page 478.

$$\rho^2 = \frac{11 \times (28,86 - 42)^2 + 15\,(39,33 - 42)^2 + 8\,(49,06 - 42)^2 + 4\,(62,5 - 42)^2 + 2\,(65 - 42)^2}{6\,377,5} = 0,773$$

Le rapport de corrélation ρ^2 n'est défini que par son carré. Dans le

cas de la corrélation linéaire, on emploiera de préférence le coefficient de corrélation *r*.

Coefficient de Spearman

Les deux paramètres que nous venons de définir : coefficient de corrélation et rapport de corrélation, ne peuvent être utilisés que lorsque les deux caractères sont quantitatifs. Dans le cas où les deux caractères sont qualitatifs, on utilise le coefficient de corrélation des rangs de Spearman.

Supposons 16 élèves, classés du point de vue de l'intelligence générale par des professeurs de lettres, d'une part, et par des professeurs de sciences, d'autre part.
On a obtenu les deux suites suivantes :

élèves	:	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
1er classement	:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2e classement	:	2	8	3	5	1	9	11	4	7	10	12	14	13	15	16	6

On se propose de mesurer la corrélation entre ces deux jugements. On utilisera le coefficient de Spearman, donné par la formule :

$$R = 1 - \frac{6 \Sigma d^2}{n (n^2 - 1)}$$

où d représente la différence entre les rangs d'un même élève dans chacune des suites et n le nombre des élèves.

$d^2 = 1^2 + 6^2 + 1^2 + 4^2 + 3^2 + 4^2 + 4^2 + 2^2 + 1^2 + 2^2 + 1^2 + 1^2 + 10^2 = 206$

$$R = 1 - \frac{6 \times 206}{16 \times 255} = 0{,}697$$

Ce coefficient dénote une nette corrélation entre les deux jugements.

Jugement sur échantillon

Dans le cas très fréquent où l'on veut mesurer la corrélation entre les deux caractères d'une même population à l'aide d'un échantillon prélevé sur cette population, on ne devra étendre à la population entière les résultats obtenus à partir de l'échantillon qu'avec certaines précautions.
Il s'agit ici d'un problème de jugement sur échantillon, que nous aborderons plus loin dans le cas de séries à une variable.

LA STATISTIQUE MATHÉMATIQUE◆

◆ Nous aborderons maintenant des notions qui font appel aux calculs des probabilités. Il est indispensable, pour les comprendre, de posséder de solides connaissances dans ce domaine. Ces notions font l'objet des statistiques mathématiques.

Nous avons défini la série statistique comme un ensemble de mesures pratiquées sur les éléments d'une population. Sous l'angle du calcul des probabilités, la série statistique est une loi de probabilité. Reprenons l'exemple de la série 1 ◆ : répartition des propriétés d'après le nombre de personnes actives.

◆ Voir page 460.

Personnes actives	1	2	3	4	5	Total
propriétés	11	15	8	4	2	40
fréquences	0,275	0,375	0,2	0,1	0,05	1

Désignons par X le nombre de personnes actives. X est une variable discrète pouvant prendre 5 valeurs : 1, 2, 3, 4, 5. Mais c'est une variable aléatoire, car toutes les valeurs n'ont pas les mêmes chances de se présenter. Il y a 11 chances sur 40 ou encore 275 chances sur 1 000 pour que X prenne la valeur 1. Il y a 15 chances sur 40 ou encore 375 chances sur 1 000 pour que X prenne la valeur 2, etc. Ces nombres : 0,275 ; 0,375 ; 0,2 ; 0,1 ; 0,05, que nous avons appelés fréquences de classe, sont les probabilités attachées aux différentes valeurs de la variable aléatoire X. L'ensemble des couples :

$$\binom{1}{0{,}275}\ \binom{2}{0{,}375}\ \binom{3}{0{,}2}\ \binom{4}{0{,}1}\ \binom{5}{0{,}05}$$

est la loi de probabilité de X. Ce que nous avons appelé moyenne de la série et désigné par $\bar{x}$ est l'espérance mathématique de X. En effet :

$$\bar{x} = \frac{1 \times 11 + 2 \times 15 + 3 \times 8 + 4 \times 4 + 5 \times 2}{40} = 2{,}02$$

ou encore :

$$\bar{x} = 1 \times \frac{11}{40} + 2 \times \frac{15}{40} + 3 \times \frac{8}{40} + 4 \times \frac{4}{40} + 5 \times \frac{2}{40} = 2{,}02.$$

$$= 1 \times 0{,}275 + 2 \times 0{,}375 + 3 \times 0{,}2 + 4 \times 0{,}1 + 5 \times 0{,}05$$

$$= \mathrm{E}(\mathrm{X}) = \sum_{1}^{5} \mathrm{X}_i . \Pr(\mathrm{X}_i) = 2{,}02.$$

Nous retrouvons le mode de calcul de l'espérance mathématique d'une variable aléatoire. De même, la variance σ^2 de la série statistique est la variance σ_X^2 de la variable aléatoire X :

$$\sigma^2 = \frac{(1 - 2{,}02)^2 \times 11 + (2 - 2{,}02)^2 \times 15 + (3 - 2{,}02)^2 \times 8 + (4 - 2{,}02)^2 \times 4 + (5 - 2{,}02)^2 \times 2}{40}.$$

X_i désignant les valeurs des classes :

$$\sigma^2 = \frac{\Sigma (X_i - \bar{x})^2 n_i}{40} = \Sigma (X_i - \bar{x})^2 \frac{n_i}{40}$$

et $\sigma_X^2 = \Sigma (X_i - E(X))^2 \Pr(X_i)$,

ce qui est la même chose ◆.

◆ Voir Probabilités.

Les séries statistiques sont donc des lois de probabilités empiriques, c'est-à-dire résultant de l'observation. Par ailleurs, nous savons qu'il existe des lois de probabilités théoriques telles que les valeurs de la variable aléatoire sont liées aux probabilités associées par une relation mathématique.

Il est souvent très utile d'ajuster les séries statistiques ou lois de probabilités empiriques à des lois théoriques dont les caractéristiques sont faciles à calculer. Dans le but de montrer l'intérêt de cet ajustement, nous allons passer en revue les principales lois de probabilités.

LA LOI BINÔMIALE

Considérons une urne contenant 2 boules blanches et 3 boules noires et envisageons l'épreuve consistant à extraire une boule de l'urne. L'extraction d'une boule blanche sera considérée comme événement favorable, l'événement contraire étant l'extraction d'une boule noire. Cette épreuve ne comprend que 2 éventualités :

— événement favorable de probabilité : $\frac{2}{5} = p$;

— événement contraire de probabilité : $\frac{3}{5} = q$;

$p + q = 1$.

Soit X une variable aléatoire définie comme suit : nombre de boules blanches sorties au cours d'un tirage. X est susceptible de prendre 2 valeurs : la valeur 1 avec la probabilité $\frac{2}{5}$ et la valeur 0 avec la probabilité $\frac{3}{5}$. La loi de probabilité de X est l'ensemble des couples :

$$\binom{1}{\frac{2}{5}} \binom{0}{\frac{3}{5}} \text{ ou, plus généralement : } \binom{1}{p} \binom{0}{q}.$$

La valeur moyenne de X est :

$$E(X) = 1 \times \frac{2}{5} + 0 \times \frac{3}{5} = \frac{2}{5} = 1 \times p + 0 \times q = p.$$

Sa variance ◆ est :

◆ Voir Probabilités.

$$\sigma_X^2 = (1 - p)^2 \times p + (0 - p)^2 \times q$$
$$= q^2 p + p^2 q = p q (p + q) = p q.$$

Son écart type est :

$$\sigma_X = \sqrt{p q}.$$

Supposons maintenant que l'on extraie 10 boules de l'urne en remettant la boule extraite dans l'urne après chaque extraction. Appelons X le nombre de boules blanches extraites. Cette nouvelle expérience consiste à répéter 10 fois l'épreuve précédente. Insistons sur le fait que, la boule extraite étant remise dans l'urne, c'est exactement la même épreuve qui se produit à chaque tirage.
X est une variable aléatoire qui peut prendre les valeurs 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Cherchons la probabilité attachée à l'une quelconque de ces valeurs : $X_k = k$ (avec $0 \leqslant k \leqslant 10$).
Si l'on extrait k boules blanches lors de ces 10 épreuves répétées, c'est que l'on a extrait $10 - k$ boules noires. Supposons que les 10 extractions soient faites dans l'ordre suivant :

$$\underbrace{B\ B\ B \ldots B}_{k} \qquad \underbrace{N\ N \ldots N}_{10 - k}$$

La probabilité d'un tel événement est égale au produit des probabilités relatives à chacun des 10 événements composants (ceux-ci étant indépendants), soit :

$$\left(\frac{2}{5}\right)^k \times \left(\frac{3}{5}\right)^{10-k}$$

Mais cet événement composé peut se produire de différentes façons selon l'ordre dans lequel les boules sont extraites. Il y a $C_{10}^k = \dfrac{10!}{k!\ (10 - k)!}$ façons de changer l'ordre d'extraction ◆. Il s'ensuit que la probabilité relative à l'événement « extraire k boules blanches » est égale à :

◆ Voir Combinaison, et Factorielle pour le signe !

$$\Pr(X_{k+1} = k) = C_{10}^k \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{10-k}$$
$$= \frac{10!}{k!\,(10 - k)!} \left(\frac{2}{5}\right)^k \left(\frac{3}{5}\right)^{10-k}$$

La loi de probabilité de X s'exprime donc, dans le cas général de n extractions :

$$\Pr(X_{k+1} = k) = C_n^k p^k q^{n-k}, \quad k \in \mathbb{N}, 0 \leqslant k \leqslant n \text{◆}.$$

◆ *Lire :* k appartenant à l'ensemble des entiers naturels supérieur ou égal à zéro inférieur ou égal à n.

On remarque que la somme des probabilités relatives aux différentes valeurs de X est égale à :

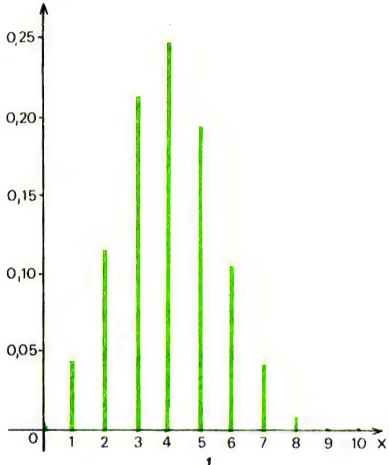
$C_n^0 q^n + C_n^1 pq^{n-1} + C_n^2 p^2 q^{n-2} + \ldots + C_n^{n-1} p^{n-1} q + C_n^n p^n = 1.$

On retrouve les termes du développement du binôme de Newton $(p + q)^n = 1$, d'où vient le nom de la loi.

Loi de probabilité de X

X_i	Pr (X_i)
0	$\left(\frac{6}{10}\right)^{10} = 0{,}006$
1	$C_{10}^1 \left(\frac{4}{10}\right) \left(\frac{6}{10}\right)^9 = 0{,}04$
2	$C_{10}^2 \left(\frac{4}{10}\right)^2 \left(\frac{6}{10}\right)^8 = 0{,}1209$
3	$C_{10}^3 \left(\frac{4}{10}\right)^3 \left(\frac{6}{10}\right)^7 = 0{,}2150$
4	$C_{10}^4 \left(\frac{4}{10}\right)^4 \left(\frac{6}{10}\right)^6 = 0{,}2508$
5	$C_{10}^5 \left(\frac{4}{10}\right)^5 \left(\frac{6}{10}\right)^5 = 0{,}2007$
6	$C_{10}^6 \left(\frac{4}{10}\right)^6 \left(\frac{6}{10}\right)^4 = 0{,}1115$
7	$C_{10}^7 \left(\frac{4}{10}\right)^7 \left(\frac{6}{10}\right)^3 = 0{,}0425$
8	$C_{10}^8 \left(\frac{4}{10}\right)^8 \left(\frac{6}{10}\right)^2 = 0{,}0106$
9	$C_{10}^9 \left(\frac{4}{10}\right)^9 \left(\frac{6}{10}\right) = 0{,}0016$
10	$C_{10}^{10} \left(\frac{4}{10}\right)^{10} = 0{,}0001$

Cette loi de probabilité se représente par un diagramme en bâtons :



Ce diagramme n'est pas symétrique, sauf si $p = q = \frac{1}{2}$. Cependant, si n augmente, il tend vers la symétrie autour de la valeur moyenne de X, et seules les valeurs de X proches de la moyenne ont une probabilité appréciable.

Valeur moyenne de X

La valeur moyenne de X peut s'évaluer par un raisonnement simple en remarquant que X est égale à la somme de 10 variables aléatoires :

x_i, x_i étant le nombre de boules blanches sorties au i-ième tirage. Par exemple, s'il sort 4 boules blanches :

$$X_5 = 4 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10}$$
$$= 1 + 1 + 1 + 1 + 0 + 0 + 0 + 0 + 0 + 0$$

en admettant cet ordre de sortie. Nous savons que ◆ :

◆ Voir Probabilités.

$$E(X) = \sum_{1}^{10} E(x_i)$$

et, comme $E(x_i) = p$: $E(X) = 10\,p$

ou, plus généralement, dans le cas de n extractions : $E(X) = n\,p$.

Les variables x_i sont indépendantes, puisqu'il s'agit d'un tirage avec remise. La boule extraite étant remise dans l'urne après chaque extraction, la composition de cette dernière est toujours la même, et les probabilités relatives à chaque épreuve ne sont pas modifiées par l'épreuve précédente. On peut donc facilement calculer la variance de X :

$$\sigma_X^2 = \Sigma\, \sigma_{x_i}^2 = npq$$

et son écart type :

$$\sigma_X = \sqrt{npq}.$$

La loi de probabilité expérimentale

La loi binômiale est une loi théorique qui peut servir de modèle à des cas concrets, chaque fois que ceux-ci peuvent être assimilés à un tirage non exhaustif dans une urne ne contenant que 2 couleurs de boule, dans les proportions p et q.
Généralement, l'urne dans laquelle on veut pratiquer les tirages est de composition inconnue et le problème consiste précisément à trouver cette composition.

Soit l'exemple suivant : une population de N électeurs comprend P d'entre eux favorables à un certain candidat et Q qui ne le sont pas. On se propose de déterminer la proportion $\frac{P}{N} = p$ d'électeurs favorables au candidat ◆. Soit une urne contenant N boules dont P sont blanches et Q noires. Peut-on déterminer le rapport $\frac{P}{N} = p$?

◆ Nous exposons ici la façon d'ajuster une série d'observations à une loi binômiale. Il ne s'agit pas de la méthode des sondages, dont le principe sera expliqué plus loin.

Extrayons 10 boules de l'urne (avec remise). On comprend facilement que cela revient à interroger 10 électeurs (pris au hasard, le même électeur pouvant être interrogé plusieurs fois).
Soit X la variable aléatoire : nombre de boules blanches sorties.

Cette variable obéit à une loi binômiale. Elle peut prendre les valeurs 0, 1, 2, ..., 10 avec les probabilités : p^{10}, $C_{10}^{1} pq^9$, etc., mais nous ne connaissons pas p. Pour en obtenir une estimation, nous répéterons 100 fois la série de 10 extractions en notant pour chaque série le nombre de boules blanches sorties. Nous obtenons la série statistique suivante, qui est la loi de probabilité expérimentale suivie par la variable aléatoire X.

Nombre de boules blanches	0	1	2	3	4	5	6	7	8	9	10	total
Nombre de séries n_i	2	5	12	21	25	20	10	4	1	0	0	100
fréquences f_i	0,02	0,05	0,12	0,21	0,25	0,20	0,10	0,04	0,01			1

Les fréquences sont les probabilités observées attachées aux différentes valeurs de la variable aléatoire X. Celle-ci obéit à une loi binômiale. Sa valeur moyenne est :

$$E(X) = np = 10\,p \text{ } (p \text{ étant inconnue}).$$

On obtiendra une estimation de cette valeur moyenne en calculant la moyenne de la série statistique :

$$\overline{X} = 0 \times 0{,}02 + 1 \times 0{,}05 + 2 \times 0{,}12 + 3 \times 0{,}21 + 4 \times 0{,}25 + 5 \times 0{,}20 + 6 \times 0{,}10 + 7 \times 0{,}04 + 8 \times 0{,}01 = 4.$$

Nous poserons $10\,p = 4$, donc $p = 0{,}4$.
0,4 est une estimation de la valeur de p. Il convient maintenant de vérifier si cette estimation est acceptable. Pour cela, on rapprochera les probabilités observées des probabilités théoriques, en calculant les probabilités attachées aux différentes valeurs d'une variable aléatoire, obéissant à une loi binômiale telle que $n = 10$ et $p = 0{,}4$ ◆.

◆ Il existe des tables donnant ces probabilités.

X	0	1	2	3	4	5	6	7	8	9	10
probabilités observées	0,02	0,05	0,12	0,21	0,25	0,20	0,10	0,04	0,01	0	0
probabilités théoriques	0,006	0,0403	0,209	0,215	0,251	0,201	0,111	0,0425	0,0106	0,0016	0,000

La correspondance entre les deux séries étant excellente, on est tout à fait en droit de penser que 4 électeurs sur 10 sont favorables à notre candidat.
Nous verrons plus loin qu'il existe une façon plus précise de tester les différences entre les probabilités théoriques et les probabilités observées ◆.

◆ On rencontre très souvent dans la pratique ce problème d'ajustement d'une distribution empirique à une loi binômiale. Citons quelques exemples : détermination du pourcentage d'individus possédant un certain groupe sanguin, détermination du pourcentage d'oranges avariées dans une livraison importante, appréciation du pourcentage de pièces défectueuses dans une certaine fabrication, etc.

LOI HYPERGÉOMÉTRIQUE

Prenons l'exemple de l'appréciation du pourcentage des pièces défectueuses dans une certaine fabrication et supposons que le contrôle effectué sur un ensemble d'objets fabriqués détruise les

objets. Il nous sera impossible de nous placer dans les conditions de la loi binômiale.
Soient N objets à contrôler, dont P sont bons et Q sont défectueux. On considère une urne contenant N boules dont P sont blanches et Q sont noires. On extrait encore, par exemple, 10 boules de l'urne, mais sans remettre la boule extraite dans l'urne, puisque l'extraction est sensée détruire la boule. Nous ne pouvons plus considérer que l'on répète 10 fois la même épreuve, la composition de l'urne changeant constamment. A chaque extraction correspond une épreuve différente, puisque la probabilité de sortir une boule blanche n'est plus la même.
Ce tirage sans remise est appelé tirage exhaustif. La loi s'appliquant à ce tirage est la loi hypergéométrique.

On ajuste assez rarement une série empirique à une loi hypergéométrique. En effet, en général n est petit par rapport à N et, par suite, le fait d'extraire une boule modifie peu la composition de l'urne et la probabilité d'extraire une boule blanche reste sensiblement constante. On considère que, lorsque $\frac{n}{N} < \frac{1}{10}$, on peut assimiler le tirage exhaustif à un tirage non exhaustif et utiliser la loi binômiale. Dans le cas particulier où N est petit (par exemple : 10 à 20 éléments), on fait en général un inventaire complet de la population, à moins toutefois que cet inventaire ne détruise la population. C'est le seul cas où l'on sera obligé d'utiliser la loi hypergéométrique.

LOI DE POISSON

La loi de Poisson, par contre, est très souvent utilisée. Elle peut être considérée comme un cas limite de la loi binômiale.
Supposons une urne contenant 1 000 boules, dont 1 est blanche et 999 sont noires. La probabilité d'extraire 1 boule blanche est $\frac{1}{1\,000} = 0{,}001$. Elle est très faible. Si l'on extrait n boules de l'urne (avec remise), il faudra répéter cette épreuve très souvent pour avoir quelque chance de sortir la boule blanche. Donc, n sera très grand ; par exemple, 5 000 tirages. Le produit $n\,p$ vaut $5\,000 \times 0{,}001 = 5$.
Dans ce cas particulier de la loi binômiale : p très petit et n suffisamment grand pour que le produit $n\,p$ soit un nombre de l'ordre de quelques unités, la loi de Poisson permettra de calculer plus simplement les résultats.
Remarquons dès maintenant que, si X est le nombre de boules blanches sorties et si l'on fait n tirages, les valeurs que peut prendre la variable X sont : 0, 1, 2, 3, ..., n ; cependant, seules les petites valeurs de X ont quelque chance de se réaliser.

Formulation de la loi de Poisson

Pour obtenir l'expression de la loi de Poisson, on part de la loi binômiale en considérant que p est petit, que n tend vers l'infini et que le produit $n\,p$ est fini.
On pose $n\,p = \lambda$. On obtient :

$$\Pr(X = k) = \frac{\lambda^k}{k!}\, e^{-\lambda}$$

où $e = 2{,}718\,28$ (base des logarithmes népériens).

Cette expression de la probabilité ne dépend que d'un paramètre k. Il existe des tables donnant la valeur des différentes probabilités des valeurs d'une variable aléatoire distribuée selon une loi de Poisson. Nous en donnons un petit extrait.

k \ λ	0,5	1	2	3	4	5
0	0,606	0,368	0,135	0,050	0,018	0,007
1	0,304	0,368	0,271	0,149	0,073	0,034
2	0,076	0,184	0,271	0,224	0,147	0,085
3	0,013	0,061	0,181	0,224	0,196	0,141
4	0,001	0,015	0,090	0,168	0,195	0,176
5		0,003	0,036	0,101	0,156	0,176
6		0,001	0,012	0,050	0,104	0,147
7			0,003	0,022	0,060	0,104
8			0,001	0,008	0,030	0,064
9				0,003	0,013	0,036
10				0,001	0,005	0,018
11					0,002	0,008
12					0,001	0,003
13						0,001

La valeur moyenne de X est :

$$E(X) = \lambda$$

La variance de X est :

$$\sigma_X^2 = \lambda$$

L'écart type de X est :

$$\sigma_X = \sqrt{\lambda}$$

La loi de Poisson trouve une application dans les phénomènes de passages aléatoires, tels : le nombre de clients entrant dans un

magasin par minute, le nombre de véhicules passant en un certain point de l'autoroute par minute, le nombre d'avions atterrissant dans un aéroport par minute, le nombre d'accidents mortels survenus en France par week-end, le nombre d'appels téléphoniques par minute, etc.

Ajustement à une loi de Poisson

Nous allons donner un exemple d'ajustement à une loi de Poisson. Supposons qu'un centre de renseignements veuille évaluer le nombre de coups de téléphone qu'il reçoit en une minute, de façon à être en mesure d'améliorer la capacité de ses installations téléphoniques.
Divisons le temps en fractions très petites, que nous appellerons temps élémentaires. Il y en a de deux sortes : les temps où il se produit un appel et ceux où il ne se produit rien. Considérons une population très grande de N temps élémentaires. Il y en a P d'entre eux où se produit un appel et Q où il ne s'en produit pas. Le temps élémentaire étant très petit (quelques millièmes de seconde), le nombre P est petit par rapport à N.
Cela correspond à une urne où il y a N boules dont P sont blanches et Q noires. $\frac{P}{N} = p$ est la probabilité d'extraire une boule blanche.
Pendant une minute, il y a n temps élémentaires. On remarquera par la suite qu'il est inutile de connaître la valeur de n ; il suffit de supposer qu'il est très grand. Compter le nombre d'appels pendant une minute revient à extraire n boules de l'urne en les remettant après chaque extraction, puisque chaque appel est indépendant l'un de l'autre. La variable aléatoire X : nombre de boules blanches sorties ou nombre d'appels en une minute, peut prendre les valeurs : 0, 1, 2, 3, ..., n, mais, p étant très petit, seules les petites valeurs de X auront une probabilité appréciable. Nous nous proposons de connaître les différentes probabilités attachées aux valeurs de X. Pour cela, nous allons faire une évaluation expérimentale, en comptant le nombre d'appels pendant une minute et en répétant cette expérience 600 fois. Nous obtenons la série suivante (trois premières lignes du tableau).

Nombre d'appels	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
nombre de minutes	17	41	90	116	118	96	62	35	11	9	4	1	0	0
fréquences	0,028	0,068	0,150	0,193	0,196	0,160	0,103	0,058	0,018	0,015	0,006	0,001	0	0
probabilités pour $\lambda = 4$	0,018	0,073	0,147	0,196	0,195	0,156	0,104	0,060	0,030	0,013	0,005	0,002	0,001	
probabilités cumulées	0,018	0,091	0,238	0,434	0,629	0,785	0,889	0,949	0,979	0,992	0,997	0,999	1	

La valeur moyenne de cette série est égale à :

$$0 \times 0{,}028 + 1 \times 0{,}068 + 2 \times 0{,}15 + 3 \times 0{,}193 + 4 \times 0{,}196 + 5 \times 0{,}16 + 6 \times 0{,}103 + 7 \times 0{,}058 + 8 \times 0{,}018 + 9 \times 0{,}015 + 10 \times 0{,}006 + 11 \times 0{,}001 = 3{,}915.$$

Cette valeur moyenne donne une évaluation de λ qui est la valeur moyenne d'une variable aléatoire distribuée selon une loi de Poisson. Il faut maintenant calculer les probabilités théoriques par la formule de Poisson et les rapprocher des probabilités observées. Les résultats pour $\lambda = 4$ sont portés sur l'avant-dernière ligne du tableau ci-dessus. Le rapprochement est très satisfaisant et nous pouvons très bien admettre que cette agence reçoit en moyenne 4 coups de téléphone à la minute et qu'il y a à peu près 95 chances sur 100 pour qu'il n'y ait pas plus de 7 coups de téléphone en une minute (dernière ligne).
On remarque que, pour utiliser la loi de Poisson, il suffit de connaître un seul paramètre λ. Les valeurs de n et de p n'interviennent pas directement dans le calcul des probabilités. Il suffit de savoir que n est grand, p très petit et que l'événement étudié peut être assimilé à une suite de tirages dans une urne.

LOI NORMALE

La loi normale est souvent appelée « loi du hasard », et l'on peut expliquer ce paradoxe de la façon suivante : un événement dû au hasard peut être considéré comme le résultat d'un grand nombre de causes, indépendantes entre elles, exerçant chacune une petite influence, du même ordre de grandeur. Soit X une variable aléatoire qui est égale à la somme de plusieurs autres :

$$X = x_1 + x_2 + \ldots + x_i + \ldots + x_n.$$

Les variables x_i sont indépendantes entre elles ; elles obéissent à des lois de probabilités quelconques, mais toutes de faible variance ; elles prennent des valeurs du même ordre de grandeur. Si n est assez grand, la variable X obéit à une loi normale ou loi de Laplace-Gauss. Cet ensemble de conditions (appelées conditions de Borel) se trouve souvent réalisé dans les phénomènes naturels qui sont le résultat de multiples facteurs indépendants entre eux. C'est ce qui explique l'importance du rôle joué par la loi normale dans l'étude des distributions expérimentales.

x étant une variable continue, prenant ses valeurs dans $\mathbb{R}$◆, on dit qu'elle obéit à une loi normale si sa densité de probabilité ◆ est :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\bar{x})^2}{2\sigma^2}} \blacklozenge$$

◆ Ensemble des nombres réels.

◆ Voir l'article « les Probabilités », page 427.

◆ e est la base du logarithme néperien. Voir *e* et Logarithme.

x est la valeur moyenne de x et σ son écart type. La fonction de répartition de X est :

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(x-\bar{x})^2}{2\sigma^2}} \, dx.$$

La somme des probabilités relatives à toutes les valeurs de x est égale à 1 :

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\frac{-(x-\bar{x})^2}{2\sigma^2}} \, dx = 1.$$

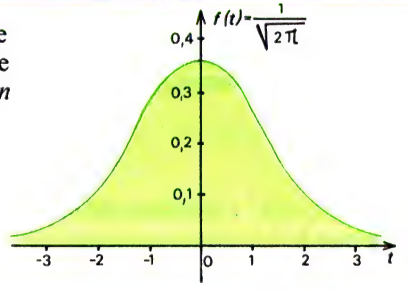
La fonction de répartition dépend de deux paramètres $\bar{x}$ et σ ; la tabulation n'est pas facile. On est conduit à considérer la variable aléatoire : $t = \dfrac{x - \bar{x}}{\sigma}$. La variable t se nomme variable centrée réduite ; sa densité de probabilité est :

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

Sa fonction de répartition est :

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{t^2}{2}} \, dt.$$

La représentation graphique de $f(t)$ est la courbe bien connue appelée bien souvent *courbe en cloche de Gauss.*



La valeur moyenne de t est nulle :

$$E(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} te^{-\frac{t^2}{2}} \, dt = 0.$$

Sa variance σ^2 est égale à 1 :

$$\sigma^2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} \, dt = 1.$$

D'où : $\sigma = 1$.

Cas limite de la loi binômiale :
la loi normale

Pour une plus grande compréhension de l'utilisation de la loi normale, nous allons étudier cette dernière en la considérant comme une approximation de la loi binômiale lorsque n est grand et p pas trop voisin de 0 ou de 1.

Nous avons déjà vu que la variable X : nombre de fois où sort la boule blanche lors de n tirages dans une urne, peut être considérée comme la somme de n variables aléatoires : $x_1, x_2, \ldots, x_n$, x_i étant le nombre de boules sorties au i-ième tirage ; nous retrouvons les conditions de Borel ◆. Si n est grand, la loi de probabilité de X tend vers une loi normale.

◆ Voir ci-dessus, page 494.

Reprenons l'exemple étudié précédemment : tirage non exhaustif de 10 boules dans une urne contenant 2 boules blanches et 3 boules noires. Soit X le nombre de boules blanches sorties. La valeur moyenne de X est :

$$np = 10 \times 0{,}4 = 4.$$

Son écart type est :

$$\sqrt{npq} = \sqrt{10 \times \frac{4 \times 6}{100}} = \frac{2}{10}\sqrt{60} = 1{,}55.$$

Nous appellerons écart réduit pour une variable aléatoire quelconque X, de valeur moyenne E (X) et d'écart type σ_X, la valeur :

$$t = \frac{X - E(X)}{\sigma_X}.$$

L'écart réduit est une nouvelle variable aléatoire qui obéit à la même loi de probabilité que X. La valeur moyenne de t est E $(t) = 0$, puisque la somme des écarts d'une variable à sa moyenne est nulle. Sa variance σ_t^2 est égale à :

$$\Sigma \left(\frac{X - E(X)}{\sigma_X} - 0\right)^2 \Pr(X) = 1.$$

Dans notre exemple :

$$t = \frac{X - 4}{1{,}55}.$$

La loi de probabilité de t est la suivante (trois premières colonnes du tableau ci-contre).

X	t	Pr (t)	Pr $(t)\sqrt{npq}$
0	$\frac{-4}{1,55} = -2,58$	0,0060	0,009
1	$\frac{-3}{1,55} = -1,93$	0,0403	0,062
2	$\frac{-2}{1,55} = -1,29$	0,1209	0,187
3	$\frac{-1}{1,55} = -0,645$	0,2150	0,333
4	0	0,2508	0,388
5	$\frac{1}{1,55} = 0,645$	0,2007	0,310
6	$\frac{2}{1,55} = 1,29$	0,1115	0,172
7	$\frac{3}{1,55} = 1,93$	0,0425	0,065
8	$\frac{4}{1,55} = 2,58$	0,0106	0,015
9	$\frac{5}{1,55} = 3,225$	0,0016	0,002
10	$\frac{6}{1,55} = 3,87$	0,0001	0,001

Bien que t soit une variable discontinue, nous représenterons la loi de probabilité de t par un histogramme en portant en abscisse les valeurs de t, les probabilités étant représentées par des rectangles. Les rectangles ont pour base un segment d'abscisse dont le centre est l'écart réduit correspondant. La surface d'un rectangle est proportionnelle à la probabilité qu'il représente.

Calculons la base d'un rectangle, en remarquant que les bases sont toutes égales. Soient 2 valeurs consécutives quelconques de t :

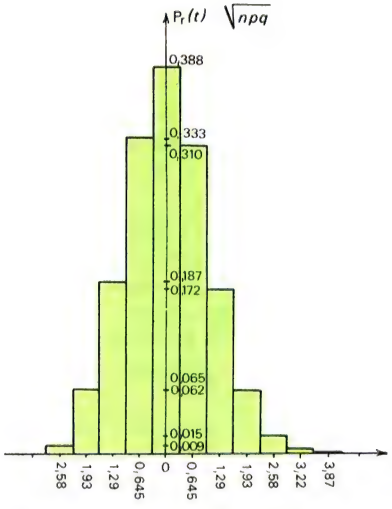
$$\frac{k+1-np}{\sqrt{npq}} - \frac{k-np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}} = \frac{1}{1,55} = 0,645.$$

La hauteur d'un rectangle est donc proportionnelle à :

$$\frac{\Pr\ (t)}{\frac{1}{\sqrt{npq}}} = \Pr\ (t)\sqrt{npq}.$$

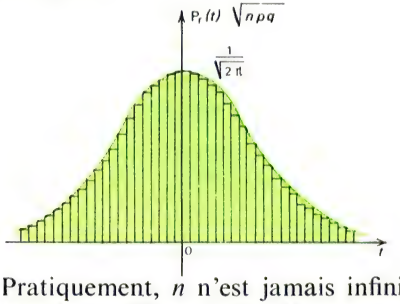
Voir les calculs dans la quatrième colonne du tableau ci-dessus.
La somme des surfaces de tous les rectangles est égale à 1. Que se passe-t-il lorsque n augmente et tend vers l'infini ? La base de chaque rectangle, qui est égale à $\sqrt{npq}$, tend vers 0, et, par suite, la probabilité d'un écart donné tend vers 0. La base de l'histogramme, qui est égale à :

$$\frac{n+1}{\sqrt{npq}},$$ tend vers l'infini.



L'histogramme est de plus en plus symétrique par rapport à l'axe des ordonnées. Nous l'avons déjà signalé en étudiant la loi binômiale. Seules les valeurs de t voisines de sa moyenne 0 ont une probabilité appréciable.

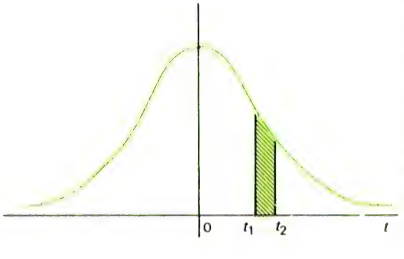
Le contour polygonal tend vers une courbe continue symétrique par rapport à l'axe des ordonnées et asymptote à l'axe des x. Son ordonnée maximale est $\frac{1}{\sqrt{2\pi}}$ et la surface entre la courbe et l'axe des abscisses est égale à 1.



On peut établir l'équation de cette courbe en cherchant la limite de la hauteur du rectangle correspondant à une valeur quelconque de t lorsque n tend vers l'infini.
On obtient :

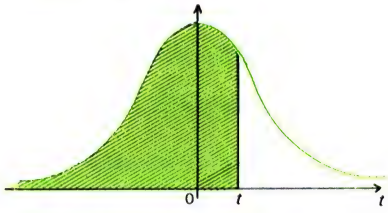
$$\Pr(t)\sqrt{npq} \simeq \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Pratiquement, n n'est jamais infini, mais il suffit que npq soit plus grand que 20 pour que l'approximation d'une loi binômiale par la loi normale soit satisfaisante.



La probabilité correspondant à un écart réduit donné est nulle puisque le rectangle correspondant a une base nulle. On cherche la probabilité pour que l'écart réduit soit compris entre deux limites t_1 et t_2. Cette probabilité est égale à la surface comprise entre la courbe, l'axe des abscisses et les deux droites $t = t_1$ et $t = t_2$.

TABLE DE GAUSS



Voici un extrait de table de Gauss :

Il existe des tables donnant les valeurs de

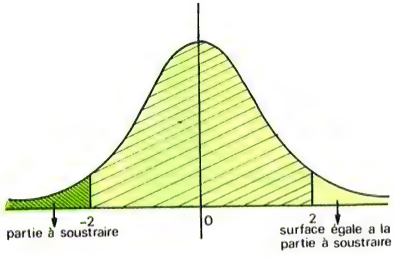
$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt,$$

soit la surface hachurée sur la figure ◆.

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{t^2}{2}} \, dt.$$

◆ La symétrie de la courbe permet de calculer n'importe quelle surface.

t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$	t	$\Phi(t)$
0,00	0,500	0,40	0,655	0,80	0,788	1,20	0,885	1,60	0,945	2,00	0,977	2,40	0,992	2,80	0,997
0,01	0,504	0,41	0,659	0,81	0,791	1,21	0,887	1,61	0,946	2,01	0,978	2,41	0,992	2,81	0,997
0,02	0,508	0,42	0,663	0,82	0,794	1,22	0,889	1,62	0,947	2,02	0,978	2,42	0,992	2,82	0,998
0,03	0,512	0,43	0,666	0,83	0,797	1,23	0,891	1,63	0,948	2,03	0,979	2,43	0,992	2,83	0,998
0,04	0,516	0,44	0,670	0,84	0,799	1,24	0,892	1,64	0,949	2,04	0,979	2,44	0,993	2,84	0,998
0,05	0,520	0,45	0,674	0,85	0,802	1,25	0,894	1,65	0,950	2,05	0,980	2,45	0,993	2,85	0,998
0,06	0,524	0,46	0,677	0,86	0,805	1,26	0,896	1,66	0,951	2,06	0,980	2,46	0,993	2,86	0,998
0,07	0,528	0,47	0,681	0,87	0,808	1,27	0,898	1,67	0,952	2,07	0,981	2,47	0,993	2,87	0,998
0,08	0,532	0,48	0,684	0,88	0,811	1,28	0,900	1,68	0,953	2,08	0,981	2,48	0,993	2,88	0,998
0,09	0,536	0,49	0,688	0,89	0,813	1,29	0,901	1,69	0,954	2,09	0,982	2,49	0,994	2,89	0,998
0,10	0,540	0,50	0,691	0,90	0,816	1,30	0,903	1,70	0,955	2,10	0,982	2,50	0,994	2,90	0,998
0,11	0,544	0,51	0,695	0,91	0,819	1,31	0,905	1,71	0,956	2,11	0,983	2,51	0,994	2,91	0,998
0,12	0,548	0,52	0,698	0,92	0,821	1,32	0,907	1,72	0,957	2,12	0;983	2,52	0,994	2,92	0,998
0,13	0,552	0,53	0,702	0,93	0,824	1,33	0;908	1,73	0,958	2,13	0,983	2,53	0,994	2,93	0,998
0,14	0,556	0,54	0,705	0,94	0,826	1,34	0,910	1,74	0,959	2,14	0,984	2,54	0,994	2,94	0,998
0,15	0,560	0,55	0,709	0,95	0,829	1,35	0,911	1,75	0,960	2,15	0,984	2,55	0,995	2,95	0,998
0,16	0,564	0,56	0,712	0,96	0,831	1,36	0,913	1,76	0,961	2,16	0,985	2,56	0,995	2,96	0,998
0,17	0,567	0,57	0,716	0,97	0,834	1,37	0,915	1,77	0,962	2,17	0,985	2,57	0,995	2,97	0,998
0,18	0,571	0,58	0,719	0,98	0,836	1,38	0;916	1,78	0,962	2,18	0,985	2,58	0,995	2,98	0,999
0,19	0,575	0,59	0,722	0,99	0,839	1,39	0,918	1,79	0,964	2,19	0,986	2,59	0,995	2,99	0,999
0,20	0,579	0,60	0,726	1,00	0,841	1,40	0,919	1,80	0,964	2,20	0,986	2,60	0,995	3,00	0,999
0,21	0,583	0,61	0,729	1,01	0,844	1,41	0,921	1,81	0,965	2,21	0,986	2,61	0,995		
0,22	0,587	0,62	0,732	1,02	0,846	1,42	0,922	1,82	0,966	2,22	0,987	2,62	0,996		
0,23	0,591	0,63	0,736	1,03	0,848	1,43	0,923	1,83	0,966	2,23	0,987	2,63	0,996		
0,24	0,595	0,64	0,739	1,04	0,851	1,44	0,925	1,84	0,967	2,24	0,987	2,64	0,996		
0,25	0,599	0,65	0,742	1,05	0,853	1,45	0,926	1,85	0,968	2,25	0,988	2,65	0,996		
0,26	0,603	0,66	0,745	1,06	0,855	1,46	0,928	1,86	0,969	2,26	0,988	2,66	0,996		
0,27	0,606	0,67	0,749	1,07	0,858	1,47	0,929	1,87	0,969	2,27	0,988	2,67	0,996		
0,28	0,610	0,68	0,752	1,08	0,860	1,48	0,931	1,88	0,970	2,28	0,989	2,68	0,996		
0,29	0,614	0,69	0,755	1,09	0,862	1,49	0,932	1,89	0,971	2,29	0,989	2,69	0,996		
0,30	0,618	0,70	0,758	1,10	0,864	1,50	0,933	1,90	0,971	2,30	0,989	2,70	0,996		
0,31	0,622	0,71	0,761	1,11	0,866	1,51	0,934	1,91	0,972	2,31	0,990	2,71	0,997		
0,32	0,625	0,72	0,764	1,12	0,869	1,52	0,936	1,92	0,973	2,32	0,990	2,72	0,997		
0,33	0,629	0,73	0,767	1,13	0,871	1,53	0,937	1,93	0,973	2,33	0,990	2,73	0,997		
0,34	0,633	0,74	0,770	1,14	0,873	1,54	0,938	1,94	0,974	2,34	0,990	2,74	0,997		
0,35	0,637	0,75	0,773	1,15	0,875	1,55	0,939	1,95	0,974	2,35	0,991	2,75	0,997		
0,36	0,641	0,76	0,776	1,16	0,877	1,56	0,941	1,96	0,975	2,36	0,991	2,76	0,997		
0,37	0,644	0,77	0,779	1,17	0,879	1,57	0,942	1,97	0,976	2,37	0,991	2,77	0,997		
0,38	0,648	0,78	0,782	1,18	0,881	1,58	0,943	1,98	0,976	2,38	0,991	2,78	0,997		
0,39	0,652	0,79	0,785	1,19	0,883	1,59	0,944	1,99	0,977	2,39	0,992	2,79	0,997		



On remarque sur la table, par exemple :

$$\Phi(2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{2} e^{-\frac{t^2}{2}} dt = 0{,}977$$

et, par suite :

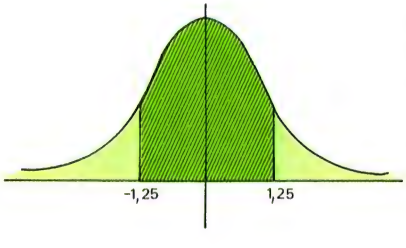
$$\frac{1}{\sqrt{2\pi}} \int_{-2}^{+2} e^{-\frac{t^2}{2}} dt$$
$$= 0{,}977 - (1 - 0{,}977)$$
$$= 0{,}954.$$

Il y a 95 chances sur 100 pour que l'écart réduit se trouve compris entre — 2 et + 2.

Utilisation de la table de Gauss

Nous allons traiter un exemple qui montrera la façon d'utiliser la table de Gauss.
Supposons qu'une entreprise fabrique des pièces dont la longueur doit être de 10 cm, avec une tolérance de $\frac{1}{100}$, c'est-à-dire que la longueur des pièces utilisables doit être comprise entre 9,99 et 10,01 cm. Par suite d'un grand nombre de facteurs liés au fonctionnement des machines, la longueur des pièces varie. Des observations ont permis d'établir la longueur moyenne des pièces, soit 10 cm, ainsi que l'écart type de la distribution des longueurs, soit 0,008 cm (nous verrons plus loin comment des paramètres sont déterminés dans la pratique). On peut se poser plusieurs questions.

a) *Quelle est la probabilité d'obtenir des pièces dont la longueur est comprise entre 9,99 et 10,01 cm ?*



La variable aléatoire X est la longueur de la pièce. Son écart réduit est égal à :

$$t = \frac{X - E(X)}{\sigma_X} = \frac{X - 10}{0{,}008}$$

Si la valeur de X varie entre 9,99 et 10,01, l'écart réduit varie entre t_1 et t_2 :

$$t_1 = \frac{9{,}99 - 10}{0{,}008} = -1{,}25 ;$$

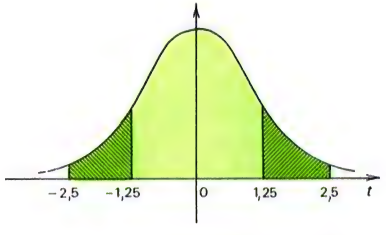
$$t_2 = \frac{10{,}01 - 10}{0{,}008} = 1{,}25.$$

Nous devons évaluer la surface sous la courbe comprise entre les deux droites $t_1 = -1{,}25$ et $t_2 = 1{,}25$. Elle est égale à :

$$\Phi(1{,}25) - (1 - \Phi(1{,}25)) = 0{,}894 - (1 - 0{,}894) = 0{,}788.$$

On a donc 78,8 chances sur 100 de produire des pièces utilisables.

b) *Quelle est la probabilité d'obtenir des pièces dont la longueur est comprise entre 9,98 cm et 9,99 cm ?*



Calculons l'écart réduit correspondant à ces deux longueurs :

$$t_3 = \frac{9{,}98 - 10}{0{,}008} = -2{,}5$$

et $t_1 = \dfrac{9{,}99 - 10}{0{,}008} = -1{,}25.$

La probabilité est égale à la surface sous la courbe comprise entre les droites $t_3 = -2{,}5$ et $t_1 = -1{,}25$.

Nous remarquons que cette surface est égale à celle qui est comprise entre les droites $t_2 = 1{,}25$ et $t_4 = 2{,}5$. Elle est donc égale à :

$$\Phi(t_4) - \Phi(t_2) = 0{,}994 - 0{,}894 = 0{,}1.$$

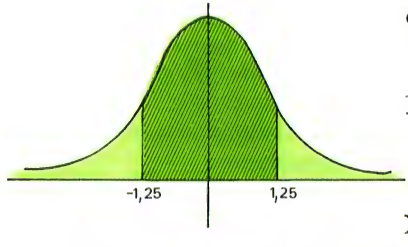
On a 1 chance sur 100 d'obtenir des pièces dont la longueur est comprise entre 9,98 cm et 9,99 cm.

c) *Entre quelles limites de part et d'autre de la moyenne se trouveront :*

1) *50 % de la fabrication ?*

2) *95 % de la fabrication ?*

1) Soient t_5 et t_6 les écarts réduits correspondant à ces limites.



On doit avoir :

$$\Phi(t_6) - (1 - \Phi(t_6)) = 0{,}50.$$

$$\Phi(t_6) = 0{,}75.$$

La table donne $t_6 = 0{,}67$, soit :

$$t_6 = \frac{X - 10}{0{,}008} = 0{,}67$$

$$X_6 = 0{,}67 \times 0{,}008 + 10 = 10{,}005$$

$$t_5 = -0{,}67 \quad X_5 = -0{,}67 \times 0{,}008 + 10 = 9{,}995.$$

50 % de la fabrication donnera des pièces dont la longueur est comprise entre 9,995 et 10,005 cm.

2) $\Phi(t_7) - (1 - \Phi(t_7)) = 0{,}95.$

$$\Phi(t_7) = 0{,}975.$$

La table donne $t_7 = 1{,}96$.

$$X_7 = 10 + 1{,}96 \times 0{,}008 = 10{,}015$$
$$X_8 = 10 - 1{,}96 \times 0{,}008 = 9{,}984$$

95 % des pièces auront une longueur comprise entre 9,984 et 10,015 cm.

Il reste bien entendu que tous ces résultats ne sont valables que dans le cas où la longueur des pièces d'une fabrication se distribue selon une loi normale. On peut supposer que c'est le cas si les erreurs sont dues à des causes multiples liées au fonctionnement de la machine et pratiquement inévitables. Mais il peut arriver que des erreurs systématiques s'ajoutent aux premières, celles-ci pouvant être attribuées au vieillissement des machines, par exemple.

AJUSTEMENT D'UNE SÉRIE STATISTIQUE À UNE LOI NORMALE

Pour déceler ces erreurs systématiques, il faut pratiquer une série d'observations et ajuster les fréquences observées à une loi normale. Nous allons donner un exemple de cet ajustement. Reprenons la production précédente, et prélevons un échantillon de 1 000 pièces dont nous mesurons la longueur. On obtient la répartition du tableau ci-contre (voir les deux premières colonnes du tableau, où les calculs sont rassemblés).

Explication des calculs :

Colonne 1 : limite des classes.

Colonne 2 : nombre de pièces ayant la longueur indiquée dans la colonne 1.

Colonne 3 : fréquences observées.

Colonne 4 : fréquences cumulées.

Colonne 5 : calcul de la longueur moyenne : somme des produits des valeurs centrales des classes par la fréquence correspondante.

Colonne 6 : calcul des écarts des valeurs centrales des classes à la moyenne.

Colonne 7 : carré de ces écarts.

1	2	3	4	5	6	7	8	9	10	11	12
longueur des pièces	nombre de pièces	fréquences observées	fréquences cumulées	produits valeur centrale × fréquences	écarts à la moyenne	carré des écarts	produit carré des écarts × fréquences	écarts réduits $\frac{X-10}{0.8}$	probabilité théorique $\Phi(t)$	probabilité théorique de chaque classe	effectifs théoriques de classe
997 à 997,5	2	0,002	0,002	1,995	-2,75	7,56	0,015	-3,12	0,001	0,001	1
997,5 à 998	7	0,007	0,009	6,984	-2,25	5,06	0,036	-2,50	0,006	0,005	5
998 à 998,5	24	0,024	0,033	23,958	-1,75	3,06	0,074	-1,87	0,031	0,025	25
998,5 à 999	80	0,08	0,113	79,9	-1,25	1,56	0,125	-1,25	0,106	0,75	75
999 à 999,25	62	0,062	0,175	61,946	-0,875	0,766	0,048	-0,937	0,174	0,068	68
999,25 à 999,50	91	0,091	0,266	90,943	-0,625	0,391	0,036	-0,625	0,266	0,092	92
999,50 à 999,75	105	0,105	0,371	104,96	-0,375	0,141	0,015	-0,312	0,378	0,112	112
999,75 à 1000	127	0,127	0,498	126,987	-0,126	0,016	0,002	0	0,5	0,122	122
1000 à 1000,25	125	0,125	0,623	125,016	0,125	0,016	0,002	0,312	0,622	0,122	122
1000,25 à 1000,50	115	0,115	0,738	115,045	0,375	0,141	0,016	0,625	0,734	0,112	112
1000,50 à 1000,75	89	0,089	0,827	89,056	0,625	0,391	0,035	0,937	0,826	0,092	92
1000,75 à 1001	67	0,067	0,894	67,059	0,875	0,766	0,051	1,25	0,894	0,068	68
1001 à 1001,5	78	0,078	0,972	78,098	1,25	1,56	0,123	1,87	0,969	0,075	75
1001,5 à 1002	20	0,02	0,992	20,04	1,75	3,06	0,061	2,50	0,994	0,025	25
1002 à 1002,5	5	0,005	0,997	5,012	2,25	5,06	0,025	3,12	0,999	0,005	5
1002,5 à 1003	3	0,003	1	3,008	2,75	7,56	0,023	3,75	1	0,001	1
	1000	1					0,687				
							$\sigma^2 = 0,687$				
							$\sigma = 0,8$				

Colonne 8 : calcul de la variance : somme des produits des carrés des écarts par la fréquence correspondante.

Colonne 9 : calcul de l'écart réduit correspondant à la limite supérieure de chaque classe.

Colonne 10 : probabilités théoriques de ces écarts, lues dans la table de Gauss.

Colonne 11 : fréquences théoriques de chaque classe.

Colonne 12 : effectifs théoriques de chaque classe.

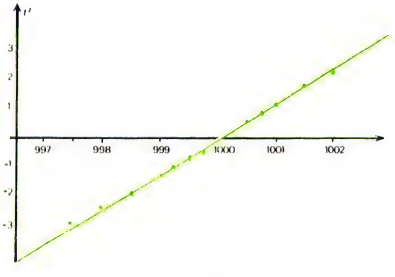
Ce procédé d'ajustement consiste à déterminer les deux paramètres : valeur moyenne et écart type, à partir de la série observée. On calcule alors l'écart réduit correspondant aux limites supérieures des classes et l'on cherche sur la table de Gauss la probabilité pour que l'écart réduit soit inférieur aux écarts calculés. Le rapprochement de ces probabilités théoriques avec les fréquences observées cumulées permet de juger si la distribution est normale. On verra ultérieurement un procédé plus précis permettant de juger de la validité de cet ajustement.

Il existe une autre méthode d'ajustement couramment employée, appelée méthode de la « droite de Henry ». Elle repose sur la relation linéaire qui existe entre la variable X et son écart réduit $t = \frac{X - E(X)}{\sigma_X}$.

La série statistique est un ensemble de classes dont les limites sont $x_0, x_1, x_2, \ldots, x_k, \ldots, x_n$ et dont les fréquences sont $f_1, f_2, \ldots, f_n$. Aux limites de classe correspondent les écarts réduits $t_0, t_1, t_2, \ldots, t_n$. La probabilité empirique pour que X prenne une valeur inférieure à x_k est égale à $\sum_1^k f_i$. La probabilité théorique est $\Phi(t_i)$.

La méthode consiste à chercher dans la table $\Phi(t)$ l'écart réduit t' correspondant aux fréquences cumulées observées. Si la variable X est distribuée selon une loi normale, les fréquences observées coïncident avec les fréquences théoriques, et l'écart réduit t' des fréquences théoriques sera égal ou très voisin de l'écart réduit observé t. Les points ayant pour coordonnées x_i et t'_i seront pratiquement alignés, puisqu'il existe une relation linéaire entre x_i et t_i. On construit ces points et l'on cherche l'équation d'une droite ajustant l'ensemble des points. La pente de la droite donnera $\frac{1}{\sigma}$ et l'ordonnée à l'origine $\frac{E(X)}{\sigma}$.

limite des classes	fréquences observées cumulées	écarts réduits correspondant aux fréquences observées
997,	0	
997,5	0,002	-2,92
998	0,009	-2,37
998,5	0,033	-1,84
999	0,113	-1,21
999,25	0,175	-0,93
999,50	0,266	-0,62
999,75	0,371	-0,33
1000	0,498	0
1000,25	0,623	0,31
1000,50	0,738	0,64
1000,75	0,827	0,94
1001	0,896	1,26
1001,50	0,972	1,91
1002	0,992	2,41
1002,50	0,997	2,75
1003	1	



Il existe de nombreux procédés d'ajustement des points à une droite. Ce problème se rencontre très souvent en statistique. Nous ne le développerons pas ici. Choisissons une façon très simple, qui consiste à trouver l'équation de la droite passant par deux points.

Soient les deux points A $\begin{pmatrix} 998,5 \\ -1,84 \end{pmatrix}$ et B $\begin{pmatrix} 1\,001 \\ 1,26 \end{pmatrix}$ l'équation de la droite AB.

$$\frac{x - 998,5}{1\,001 - 998,5} = \frac{t + 1,84}{1,26 + 1,84}$$

$$t = \frac{3,10\,x}{2,5} - 1\,239,98.$$

On trouve $\sigma = \frac{2,5}{3,10} = 0,8\,064$.

$$E(X) = 1\,239,98 \times 0,8\,064 = 999,91.$$

Nous avons passé en revue les principales lois de probabilités dont l'utilisation est très fréquente en statistique. Il en existe d'autres moins employées, telles que les distributions de Gram-Charlier, de Galton, de Pearson, dont l'étude sortirait du cadre de cet article. Cependant, les problèmes de jugement sur échantillons que nous allons aborder maintenant nous conduiront à faire appel à deux distributions particulières. Il s'agit d'une distribution de Student-Fisher ◆ et d'une distribution de Pearson. Cette dernière est également utilisée comme test d'ajustement à une loi de probabilité.

◆ La distribution de Student-Fisher est utilisée dans l'étude des petits échantillons (moins de 30 éléments).

Nous avons vu, en effet, que l'ajustement consistait à rapprocher les probabilités théoriques et les probabilités observées. Le test de Pearson permet d'évaluer la limite acceptable des différences entre les deux probabilités.

LE TEST DE PEARSON

Reprenons l'ajustement précédent relatif aux distributions des longueurs de 1 000 tubes métalliques. Posons N = 1 000. La variable aléatoire X prend les valeurs (voir le tableau page 503) :

997,25 cm, 997,75 cm, ... (centre des classes, col. 1).

Soient :

$n_1, n_2, \ldots, n_{16}$ le nombre de fois où les valeurs se rencontrent (col. 2) ;

$p_1, p_2, \ldots, p_{16}$ les probabilités théoriques attachées à ces valeurs (col. 11) ;

$Np_1, Np_2, \ldots, Np_{16}$ le nombre de fois théorique où l'on devrait rencontrer chaque valeur de X (col. 12).

$n_1 - Np_1, n_2 - Np_2, \ldots, n_{16} - Np_{16}$ sont les différences que nous nous proposons de tester.

On conçoit que plus ces différences sont grandes, moins on peut donner de confiance à l'ajustement. La somme de ces différences ne nous donnerait aucun renseignement, étant donné que leur signe est variable. On est amené ainsi à faire la somme des carrés de ces différences : $(n_i - Np_i)^2$ ou encore des expressions : $\frac{(n_i - Np_i)^2}{Np_i}$.

Soit : $\chi^2 = \frac{(n_1 - Np_1)^2}{Np_1} + \frac{(n_2 - Np_2)^2}{Np_2} + \ldots + \frac{(n_{16} - Np_{16})^2}{Np_{16}}$.

Chaque élément de la somme précédente peut être considéré comme le carré d'une variable aléatoire. Les nombres n_1, n_2, ..., n_{16} sont en effet aléatoires, puisqu'on a prélevé 1 000 tubes au hasard pour faire l'expérience. On conçoit qu'en prélevant 1 000 autres tubes on trouverait des résultats différents. Ces variables sont indépendantes. En réalité, 15 seulement d'entre elles le sont, la 16^e étant obligatoirement égale à 1 000 moins la somme des 15 premières.
La variable χ^2, qui est égale à la somme des carrés de ν variables aléatoires indépendantes obéissant toutes à la loi normale réduite, suit une loi de Pearson à $\nu - 1$ degrés de liberté (puisqu'en réalité il n'y a que $\nu - 1$ variables indépendantes). Sa densité de probabilité est :

$$\alpha e^{-\frac{\chi^2}{2}} \chi^{\nu-1} \, d\chi,$$

où α est une constante, déterminée par l'égalité :

$$\alpha = \frac{1}{\int_{-\infty}^{+\infty} e^{-\frac{\chi^2}{2}} \chi^{\nu-1} \, d\chi}.$$

Il existe des tables fournissant les valeurs correspondantes de χ^2, ν et $F(\chi^2) = \Pr(\chi^2 \geqslant x^2)$.

Extrait d'une table de distribution de χ^2

$F(\chi^2)$	0,990	0,950	0,05
1	0,00016	0,00393	3,84
2	0,02	0,10	5,99
3	0,12	0,35	7,81
4	0,30	0,71	9,49
5	0,55	1,15	11,07
6	0,87	1,64	12,59
7	1,24	2,17	14,07
8	1,65	2,73	15,51
9	2,09	3,33	16,92
10	2,56	3,94	18,31
11	3,05	4,57	19,67
12	3,57	5,23	21,03
13	4,11	5,89	22,36
14	4,66	6,57	23,68
15	5,23	7,26	25
16	5,81	7,96	26,30
17	6,41	8,67	27,59
18	7,01	9,39	28,87
19	7,63	10,1	30,14
20	8,26	10,9	31,41

Au-delà de 30 degrés de liberté, on peut remplacer la loi de Pearson par une loi normale réduite s'appliquant à la variable suivante :

$$u = \sqrt{2\chi^2} - \sqrt{2\nu - 1}.$$

Nous reprendrons l'exemple donné précédemment d'un ajustement à une loi normale pour lui appliquer le test de Pearson.

X (voir page 505)	nombre de pièces n_i	nombres théoriques Np_i	$n_i - Np_i$	$\frac{(n_i - Np_i)^2}{Np_i}$
997,25	2	1		
997,75	7	5	3	1,5
998,25	24	25	-1	0,004
998,75	80	75	5	0,33
999,125	62	68	-6	0,53
999,375	91	92	-1	0,001
999,625	105	112	-7	0,44
999,875	127	122	5	0,20
1000,125	125	122	3	0,07
1000,375	115	112	3	0,08
1000,625	89	92	-3	0,1
1000,875	67	68	1	0,01
1001,25	78	75	3	0,12
1001,75	20	25	-5	1
1002,25	5	5	2	0,66
1002,75	3	1		
				5,045

Les classes extrêmes où les effectifs sont faibles ont été groupées de façon que les effectifs théoriques des nouvelles classes soient au moins de 5.

5 est un minimum au-dessous duquel on ne doit jamais descendre. Cette condition est nécessaire à l'interprétation correcte de la valeur de χ^2 pour décider si l'on peut conserver la loi théorique ou si l'on doit en rechercher une autre. On lit dans la table, pour $\nu = 15$:

$$\Pr(\chi^2 \geqslant 5{,}23) = 0{,}99$$
$$\Pr(\chi^2 \geqslant 7{,}26) = 0{,}95$$
$$\Pr(\chi^2 \geqslant 25) = 0{,}05.$$

Plus la valeur de χ^2 augmente, plus la probabilité de rencontrer cette valeur diminue. Les fortes valeurs de χ^2 sont donc improbables et, si nous les rencontrons, il est permis de douter de la validité de notre ajustement. Par contre, dans notre exemple, $\chi^2 = 5{,}045$. Il y a plus de 99 chances sur 100 de rencontrer une valeur égale à 5,045 et nous pouvons penser que notre hypothèse est acceptable ou, tout au moins, non infirmée par ce résultat.

JUGEMENT SUR ÉCHANTILLON LES SONDAGES

Nous avons appris à ajuster des séries d'observations à des lois de probabilités théoriques de même moyenne et de même écart type que la loi expérimentale. Nous avons volontairement passé sous silence un fait très important, à savoir que les distributions empi-

riques ne portaient pas en général sur la totalité de la population, mais seulement sur une partie. Il est très rare que l'on puisse mesurer le caractère qui nous intéresse sur la population entière. En fait, dans presque tous les cas, on n'étudie qu'un échantillon. Dans quelle mesure cet échantillon est-il représentatif de la population entière? C'est le problème qui va être traité maintenant, et nous allons répondre à deux questions qui viennent naturellement à l'esprit.

TAILLE DE L'ÉCHANTILLON

Quel doit être le nombre n de prélèvements sur la population entière pour constituer un échantillon valable? On verra que la précision des observations augmente avec la taille, mais, par ailleurs, le coût de l'opération également. Il faut donc évaluer la valeur minimale de n, compatible avec le degré de précision cherché.

Nous distinguerons les « grands » échantillons et les « petits » échantillons. L'expérience prouve que, si n est supérieur à 30, l'échantillon peut être classé comme « grand » et que, si n est inférieur ou égal à 30, l'échantillon est classé dans les « petits ». Si nous désignons par m et σ la moyenne et l'écart type de la population entière, cette dernière sera correctement représentée par l'échantillon si celui-ci permet une bonne appréciation de m et de σ.

TESTS D'HYPOTHÈSES

Ayant fait le choix d'un échantillon d'une certaine taille, on a émis des hypothèses relativement à la population étudiée. Dans quelles limites l'échantillon justifie-t-il ces hypothèses? Quel degré de confiance peut-on accorder à ces hypothèses? Nous classerons ces tests en deux grandes catégories : tests d'estimation et tests de comparaison, et nous envisagerons successivement trois problèmes essentiels :
— Estimation de la moyenne vraie m et de l'écart type σ d'une population à l'aide d'un échantillon de n observations prises au hasard dans cette population.
— Comparaison de la moyenne et de l'écart type d'un échantillon à la moyenne et à l'écart type supposés connus de la population entière ◆.
— Comparaison de deux échantillons.
Supposons que nous étudions les effets de deux engrais A et B sur une récolte. Un échantillon prélevé sur la récolte A a donné une moyenne m_1, un autre échantillon, prélevé sur la récolte B, a donné une moyenne m_2. Le problème est le suivant : la différence entre m_1 et m_2 est-elle due aux fluctuations d'échantillonnage? Dans ce cas, on peut considérer que les deux échantillons ont été prélevés sur une même population (c'est-à-dire que les engrais A et B donnent

◆ C'est le cas, par exemple, d'un contrôle de fabrication. Il s'agit de vérifier si la fabrication reste conforme aux normes établies et réalisées au départ.

le même rendement et les récoltes A et B ne sont qu'une même récolte). Dans le cas contraire, il faudra conclure que les deux échantillons ont été prélevés sur des populations différentes et, par suite, que les rendements de A et de B sont différents.

SEUIL DE SIGNIFICATION

Le test d'hypothèse conduit, comme nous le verrons plus loin, au calcul d'une probabilité. Si la probabilité est très petite, on repoussera l'hypothèse, car il est peu probable qu'elle soit bonne. On dira en général que la différence est *significative*, c'est-à-dire que, la différence entre la moyenne supposée et la moyenne constatée étant peu probable, il y a peu de chance qu'elle se réalise et, par conséquent, l'hypothèse doit être fausse. Si la probabilité est grande, on dira que l'hypothèse émise est vraisemblable. Cela ne signifie pas qu'elle soit exacte, mais seulement que les différences observées peuvent être dues aux fluctuations d'échantillonnage. L'hypothèse émise n'est pas fausse à priori ; elle est plausible.
Il faut donc fixer un seuil de probabilité au-dessous duquel on rejettera systématiquement l'hypothèse, et au-dessus duquel on la considérera comme plausible. Ce seuil se nomme le *seuil de signification*. Il est important de préciser que la réponse ne sera jamais absolue. Il est toujours possible de repousser une hypothèse exacte : c'est le risque de première espèce, ou d'accepter une hypothèse fausse : c'est le risque de deuxième espèce.

LES ÉCHANTILLONS

Pour résoudre les différents problèmes que nous venons d'exposer, il faut au préalable établir quelques théorèmes fondamentaux.
Supposons que l'on veuille étudier une population sous un certain caractère, par exemple: répartition des hommes de vingt ans, de race noire, d'après leur taille. Appelons X la variable aléatoire « taille de ces individus », pouvant prendre N valeurs : x_1, x_2, x_3, ..., x_N, et soient E(X) et σ_X la taille moyenne et l'écart type de la distribution. Appelons L la loi théorique la plus proche de cette distribution.
E(X), σ_X et L sont inconnus. Or, pour les connaître, on ne peut mesurer tous les éléments de la population, mais seulement un échantillon de n individus, que nous supposerons tiré au hasard parmi la population entière comprenant N hommes.
Tout d'abord, faisons une remarque : le prélèvement de l'échantillon peut avoir lieu « avec remise » (le même individu peut figurer plusieurs fois dans l'échantillon) ou « sans remise » (un individu ne peut y figurer qu'une fois). Dans le premier cas, les variables aléatoires qui constituent l'échantillon sont indépendantes ; dans le second cas, elles ne le sont pas, ce qui entraîne une formulation plus compliquée.

On décide donc dans la pratique de considérer les échantillons comme non exhaustifs (avec remise) dans tous les cas, dès l'instant que $\frac{n}{N} < \frac{1}{10}$, c'est-à-dire : la population est très nombreuse, l'effectif de l'échantillon relativement petit par rapport à la population.
Si la population est peu nombreuse, on fera l'inventaire complet, sauf dans quelques cas où cela est impossible ; on devra alors utiliser les formules relatives aux échantillons exhaustifs.
Supposons qu'on ait prélevé un échantillon de n éléments ; mesurons la taille de chaque individu. On obtient les valeurs $x_1, x_2, \ldots, x_n$, qui sont des valeurs de la variable X. Si nous avions mesuré les N individus, chaque valeur $x_1, x_2, \ldots, x_N$ serait une valeur certaine : la taille d'un individu donné. Mais, n'en ayant mesuré que n, chaque valeur $x_1, x_2, \ldots, x_n$ est une variable aléatoire, puisqu'elle peut prendre des valeurs différentes si l'on tire successivement plusieurs échantillons. Et c'est une variable aléatoire qui est régie par la loi L de X, puisqu'elle peut prendre toutes les valeurs prises par X avec la même probabilité (aucune de ces valeurs n'est privilégiée dans l'échantillon, seule sa fréquence dans la population totale intervient pour fixer sa probabilité d'arrivée dans l'échantillon).
$\bar{x}_e$: moyenne de l'échantillon, et σ_e : écart type de l'échantillon, qui sont des combinaisons des x_i de l'échantillon, sont également des variables aléatoires. Ces variables aléatoires obéissent à une loi de probabilité et la connaissance de cette loi permettra d'établir les liens qui unissent $\bar{x}_e$ et σ_e, paramètres de l'échantillon, à E(X) et à σ_X, paramètres de la population totale.

SÉRIE DES MOYENNES

Supposons un instant que nous ayons prélevé sur la population totale tous les échantillons possibles de n éléments (prélèvements avec remise), chacun de ces échantillons ayant donné une moyenne $\bar{x}_e$ et une variance σ_e^2. Considérons alors l'ensemble de tous les $\bar{x}_e$, que nous appellerons série des moyennes, et l'ensemble de tous les σ_e^2, que nous appellerons série des variances. Proposons-nous de calculer la moyenne et la variance de la série des moyennes.
$\bar{x}_e$ est la somme de n variables indépendantes qui ont toutes même valeur moyenne, celle de la population entière, soit E(X). Nous savons que la moyenne d'une somme de variables aléatoires est égale à la somme des valeurs moyennes de chaque variable ◆ :

◆ Voir l'article « les Probabilités », pages 433, 434.

$$\bar{x}_e = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

$$E(\bar{x}_e) = \frac{1}{n}[E(X) + E(X) + \ldots + E(X)]$$

$$E(\bar{x}_e) = E(X).$$

La moyenne de la série des moyennes est la même que celle de la population entière.
Les variables aléatoires $x_1, x_2, \ldots, x_n$ étant toutes indépendantes (tirage avec remise) et ayant toutes pour variance σ_x^2, nous savons que la variance d'une somme de variables aléatoires indépendantes est égale à la somme des variances de chaque variable ◆ :

◆ Voir l'article Les probabilités, page 435.

$$\bar{x}_e = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

$$n\,\bar{x}_e = x_1 + x_2 + \ldots + x_n$$

$$\sigma^2_{n\bar{x}_e} = n\,\sigma_x^2$$

$$\sigma^2_{n\bar{x}_e} = \frac{[n\,\bar{x}_e - n\mathrm{E}(X)]^2}{n} = \frac{n^2[\bar{x}_e - \mathrm{E}(X)]^2}{n} = n^2\,\sigma^2_{\bar{x}_e}.$$

D'où : $n^2\,\sigma^2_{\bar{x}_e} = n\,\sigma_x^2$ et $\sigma^2_{\bar{x}_e} = \dfrac{\sigma_x^2}{n}$

Dans le cas où les échantillons sont prélevés sans remise, les variables $x_1, x_2, \ldots, x_n$ ne sont plus indépendantes. Nous savons que, dans ce cas, la valeur moyenne de la somme est encore égale à la somme des valeurs moyennes :

$$\mathrm{E}(\bar{x}_e) = \mathrm{E}(X).$$

Mais il n'en est pas de même des variances ; un calcul facile mais un peu long permettrait de trouver la variance $\sigma^2_{\bar{x}_e}$:

$$\sigma^2_{\bar{x}_e} = \frac{\sigma^2}{n} \times \frac{N - n}{N - 1}.$$

Si N est grand, on confond N et N — 1, ce qui donne :

$$\sigma^2_{\bar{x}_e} = \frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right).$$

On voit que, si n est petit par rapport à N, on retrouve, en négligeant $\dfrac{n}{N}$, le résultat obtenu dans le cas d'un échantillon non exhaustif.

Ainsi, donc, la série des moyennes est une distribution qui admet pour moyenne E(X) la moyenne de la population entière et pour variance $\dfrac{\sigma_x^2}{n}$ ou $\dfrac{\sigma_x^2}{n}\left(1 - \dfrac{n}{N}\right)$, selon la façon de prélever les échantillons.

SÉRIE DES VARIANCES

La série des variances est l'ensemble des variances σ_e^2 de tous les échantillons possibles. Nous nous proposons de déterminer la moyenne et la variance de la série des variances. Les calculs, quoique simples, sont un peu longs et nous nous contenterons de donner les résultats.

Moyenne de la série des variances :

$$E(\sigma_e^2) = \frac{n-1}{n}\sigma_x^2$$

Variance de la série des variances :

$$\sigma^2_{\sigma_e^2} = \frac{2}{n}\sigma^4,$$

et, par suite *écart type :*

$$\sigma_{\sigma_e^2} = \sqrt{\frac{2}{n}}\,\sigma^2$$

Il s'ensuit que, si l'on prélève un grand nombre d'échantillons, la moyenne des variances ne tend pas vers la variance de la population entière, mais vers $\frac{n-1}{n}\sigma_x^2$.

Si la taille de l'échantillon augmente, $\frac{n-1}{n}$ tend vers 1 et $E(\sigma_e^2)$ tend vers σ_x^2 en même temps que $\sigma^2_{\sigma_e^2}$ tend vers zéro ; c'est-à-dire que les valeurs de la variance tendent à se concentrer autour de sa valeur moyenne σ_x^2 lorsque la taille de l'échantillon augmente.

Lois de $\bar{x}_e$ et σ_e

Pour tirer parti de ces résultats, il nous faudrait connaître la forme de la distribution des moyennes et celle de la distribution des variances. Nous admettrons ce qui suit :

1. — Si l'on prend des échantillons au hasard dans une population normale de moyenne m et d'écart type σ, la distribution des moyennes est aussi une série normale de moyenne m et d'écart type $\frac{\sigma}{\sqrt{n}}$. De même, la distribution des variances est une série normale si $n > 30$.

2. — Si l'on prend des échantillons au hasard dans une population régie par une loi quelconque, la distribution des moyennes tend vers

une loi normale lorsque n augmente. Pratiquement, il suffit que l'on ait $n > 30$. De même pour la distribution des variances.
En utilisant ces nouvelles connaissances, nous allons reprendre les problèmes posés au début de cet article. On conçoit intuitivement que les deux problèmes : taille des échantillons et tests d'hypothèse, sont liés. Nous prendrons successivement trois exemples, à partir desquels nous répondrons aux questions qui nous occupent.

ESTIMATION D'UNE MOYENNE

On se propose de connaître la consommation moyenne annuelle d'aspirine m pour une personne adulte. Un échantillon constitué de 100 personnes a fourni une moyenne de consommation de 30 grammes avec un écart type de 10 grammes :

$$\bar{x}_e = 30 \qquad \sigma_e = 10.$$

Cette valeur moyenne est un élément de la série des moyennes : distribution normale de moyenne m, qui est précisément la valeur que nous cherchons, et d'écart type $\frac{\sigma}{\sqrt{n}}$. Rappelons que m et σ sont les paramètres de la population totale des personnes adultes. La valeur de la variable réduite correspondant à $\bar{x}_e = 30$ est :

$$t_e = \frac{\bar{x}_e - m}{\frac{\sigma}{\sqrt{n}}}.$$

Nous savons qu'il y a une probabilité de 95 % pour que t_e soit compris entre + 1,96 et — 1,96 (voir table de Gauss, page 499) :

$$\Pr(-1{,}96 \leqslant t \leqslant 1{,}96) = 0{,}95.$$

Soit :

$$\Pr\left(-1{,}96 \leqslant \frac{\bar{x}_e - m}{\frac{\sigma}{\sqrt{n}}} \leqslant 1{,}96\right) = 0{,}95$$

Ce qui se transforme facilement de la manière suivante :

$$\Pr\left(\bar{x}_e - 1{,}96\,\frac{\sigma}{\sqrt{n}} \leqslant m \leqslant \bar{x}_e + 1{,}96\,\frac{\sigma}{\sqrt{n}}\right) = 0{,}95$$

Il y a donc 95 % de chances pour que la moyenne cherchée soit comprise entre $30 - 1{,}96\,\frac{\sigma}{\sqrt{n}}$ et $30 + 1{,}96\,\frac{\sigma}{\sqrt{n}}$. Pour faire ce calcul,

il faudrait connaître σ, l'écart type de la population entière. Dans la plupart des cas, σ est inconnu ; mais nous savons que la valeur moyenne de la série des variances est $\frac{n-1}{n}\sigma^2$. On confondra σ_e^2, la variance de l'échantillon, avec sa valeur moyenne :

$$\sigma_e^2 = \frac{n-1}{n}\sigma^2.$$

Et l'on prendra comme valeur de σ :

$$\sigma = \sqrt{\frac{n}{n-1}}\,\sigma_e.$$

On peut alors calculer les limites entre lesquelles se trouve la moyenne cherchée, avec une probabilité de 95 % :

$$\bar{x}_e - \frac{1{,}96\,\sigma}{\sqrt{n}} = 30 - 1{,}96\,\frac{10}{\sqrt{99}} = 28{,}04.$$

$$\bar{x}_e + \frac{1{,}96\,\sigma}{\sqrt{n}} = 30 + 1{,}96\,\frac{10}{\sqrt{99}} = 31{,}96.$$

Il y a 95 % de chances pour que la consommation annuelle d'aspirine par un adulte varie entre 28,04 g et 31,96 g.

Risque de première espèce

La moyenne vraie m pourrait tomber à l'extérieur de ces limites avec une probabilité de 5 %. La probabilité de rejeter une hypothèse alors qu'elle est vraie est donc de 5 %. C'est le risque de première espèce ; il est égal au seuil de signification.

Risque de deuxième espèce

Le risque de deuxième espèce : accepter une hypothèse alors qu'elle est fausse, est plus difficile à évaluer. Supposons que nous décidions d'adopter la moyenne de 30 g, alors qu'en réalité elle est de 29 g. Aux limites déterminées plus haut (28,04 et 31,96) correspondent les écarts réduits :

$$\frac{28{,}04 - 29}{\frac{10}{\sqrt{99}}} = -0{,}96 \quad \text{et} \quad \frac{31{,}96 - 29}{\frac{10}{\sqrt{99}}} = 2{,}96$$

La table de Gauss donne la probabilité pour que l'écart réduit soit compris entre — 0,96 et 2,96, soit 0,829.

On voit donc que, si la moyenne est de 29, l'hypothèse sur la moyenne comprise entre 28,04 et 31,96 peut être considérée comme bonne avec une probabilité de 0,829. On a donc la probabilité 0,829 d'accepter une moyenne de 30 alors qu'elle est fausse. C'est le risque de deuxième espèce. Ce risque de deuxième espèce est difficile à évaluer, surtout dans un problème d'estimation de moyenne, puisqu'il faut connaître la moyenne exacte. Il peut pourtant être très important si, par exemple, on cherche la dose d'un produit toxique à ne pas dépasser. On peut diminuer le risque de deuxième espèce en élevant le seuil de signification, ce qui, n restant le même, va diminuer l'intervalle de confiance et, par conséquent, les possibilités d'accepter une moyenne fausse. Mais on augmente alors le risque de première espèce, puisqu'il est égal au seuil de signification.
Dans un jugement, on est toujours forcément partagé entre les deux risques.

TAILLE DE L'ÉCHANTILLON

Reprenons le problème précédent : évaluation de la consommation moyenne d'aspirine, en prélevant un échantillon de 1 000 personnes. Supposons que la moyenne calculée sur cet échantillon soit encore de 30 g et l'écart type de 10.
Il y a 95 % de chances pour que la moyenne vraie soit comprise entre :

$$30 - \frac{1{,}96 \times 10}{\sqrt{999}} = 29{,}38 \quad \text{et} \quad 30 + \frac{1{,}96 \times 10}{\sqrt{999}} = 30{,}62.$$

Les résultats sont beaucoup plus précis, comme on pouvait le prévoir. Si donc on se fixe un seuil de signification, l'intervalle de confiance varie en fonction de la taille de l'échantillon. On peut aussi déterminer la valeur de n pour que, le seuil de signification étant donné, on trouve la moyenne vraie dans un intervalle que l'on se fixe à l'avance.

ESTIMATION DE p DANS UNE LOI BINOMIALE

Nous reprendrons l'exemple qui a servi à étudier l'ajustement d'une série statistique à une loi binômiale.
Soit une population d'électeurs parmi lesquels une proportion p est favorable à un candidat. Nous voulons déterminer p en prélevant un échantillon de 1 000 éléments sur la population entière. Soit X la variable aléatoire susceptible de prendre la valeur 1 si la personne est favorable au candidat et la valeur 0 dans le cas contraire :

$$\Pr(x = 1) = p$$
$$\Pr(x = 0) = q = 1 - p$$

Nous savons que cette variable a pour valeur moyenne p et pour écart type $\sqrt{pq}$ ◆. Si nous interrogeons 1 000 personnes, nous obtiendrons 1 000 valeurs de X, et, pour moyenne de ces 1 000 valeurs :

◆ Voir page 487.

$$\bar{x}_e = \frac{x_1 + x_2 + \ldots + x_{1000}}{1\,000} = p'.$$

Soit, par exemple :

$$\bar{x}_e = \frac{390}{1\,000} = 0{,}39.$$

$\bar{x}_e$ est alors la proportion d'électeurs favorables dans l'échantillon, puisque les valeurs de $x_1, x_2, \ldots, x_{1000}$ sont égales soit à 0, soit à 1. $\bar{x}_e$ appartient à la série des moyennes dont la valeur est p et l'écart type $\frac{\sqrt{pq}}{\sqrt{n}}$.

Comme précédemment ◆, on peut écrire :

◆ Voir page précédente.

$$\Pr\left(p' - 1{,}96\sqrt{\frac{pq}{n}} < p < p' + 1{,}96\sqrt{\frac{pq}{n}}\right) = 0{,}95.$$

Ne connaissant pas p, nous prendrons pour valeur de p : p' et pour valeur de q : $1 - p'$. On obtient :

$$\Pr(36 < p < 42) = 0{,}95$$

On estimera, avec 95 % de chances d'exactitude, que la proportion d'électeurs favorables au candidat est comprise entre 36 et 42 %.

COMPARAISON D'UNE MOYENNE A DES NORMES

Supposons une machine automatique remplissant des bouteilles avec une contenance de 100 cl. Cette machine a fonctionné normalement pendant un certain temps. Des observations nombreuses ont donné un écart type $\sigma = 1$. On veut vérifier le bon état de la machine et, pour cela, on mesure le liquide contenu dans 50 bouteilles ; la moyenne est de 99,78 cl. Que peut-on conclure de cette observation ?
Si l'état de la machine est inchangé, cette moyenne de 99,78 cl est un élément de la série des moyennes qui admet, comme on le sait, la moyenne 100 et l'écart type $\frac{1}{\sqrt{50}}$. L'écart réduit correspondant à 99,78 est :

$$\frac{99{,}78 - 100}{\frac{1}{\sqrt{50}}} = -1{,}56.$$

Si l'on admet un seuil de signification de 0,05, c'est-à-dire si on considère comme anormal un écart à la moyenne qui a une probabilité inférieure ou égale à 0,05 d'être dépassé en valeur absolue, on pourra dire qu'une déviation supérieure à 1,96 fois son écart type est statistiquement significative, c'est-à-dire qu'il y a des chances pour que la population de laquelle l'échantillon est extrait ne soit pas une population conforme aux normes.
Ici, l'écart réduit est de 1,56. La table de Gauss donne une probabilité de 0,118 pour que cet écart soit dépassé en valeur absolue ; il n'est donc pas significatif. Il est possible de rencontrer une moyenne égale à 99,78 sans que la machine soit détraquée.
Remarquons que, si nous obtenons le même résultat : moyenne 99,78, avec un échantillon de 100 bouteilles, l'écart réduit :

$$\frac{99{,}78 - 100}{\frac{1}{\sqrt{100}}} = 2{,}2$$

est significatif. La table de Gauss donne une probabilité égale à 0,028 de dépasser un écart de 2,2 en valeur absolue.
Intuitivement, cela semble normal, car plus la taille de l'échantillon est grande, plus la moyenne doit se rapprocher de la valeur exacte. Si elle en reste toujours aussi éloignée, il y a donc lieu de s'en inquiéter.
A une moyenne de 99,78 donnée par un échantillon de 10 000 bouteilles correspond l'écart réduit 22,8. Il est infiniment peu probable : la machine est détraquée.
Les renseignements sont d'autant plus précis que l'échantillon est plus grand : on s'en serait douté. Il est donc prudent, dans un test de ce genre, de tirer un second échantillon, portant par exemple sur 100 bouteilles. Si la moyenne de cet échantillon se rapproche de 100 et donne une valeur réduite voisine de 1,56, les résultats du premier échantillon seront confirmés.

COMPARAISON DE DEUX ÉCHANTILLONS

On désire comparer le rendement de deux engrais A et B. Les récoltes A et B ont été divisées en parcelles sur lesquelles on a prélevé deux échantillons dont les tailles sont respectivement : 80 parcelles pour la récolte A et 50 parcelles pour la récolte B. Le rendement moyen des 80 parcelles A est : $\bar{x}_A = 80$, avec un écart type $\sigma_A = 10$. Le rendement moyen des 50 parcelles B est : $\bar{x}_B = 78$, avec un écart type $\sigma_B = 9$.
Peut-on considérer ces deux échantillons comme extraits d'une même population (et, dans ce cas, les récoltes A et B seraient semblables) ? Supposons que les récoltes A et B ne forment qu'une même popu-

lation, la moyenne inconnue de cette population étant m et son écart type σ également inconnu.
$\bar{x}_A$ et $\bar{x}_B$ sont deux éléments de la série des moyennes, dont la valeur moyenne est m.
La différence $d = \bar{x}_A - \bar{x}_B$ est une variable aléatoire qui suit une loi normale. La valeur moyenne de d est celle de $\bar{x}_A - \bar{x}_B$, soit $m - m = 0$.
$\bar{x}_A$ et $\bar{x}_B$ étant deux variables indépendantes, la variance de d est égale à $\sigma^2_{\bar{x}_A} + \sigma^2_{\bar{x}_B}$. Mais:

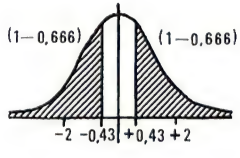
$$\sigma^2_{\bar{x}_A} = \frac{\sigma^2}{80} \text{ et } \sigma^2_{\bar{x}_B} = \frac{\sigma^2}{50}.$$

σ étant inconnu, nous remplaçons $\frac{\sigma^2}{80}$ par $\frac{\sigma^2_A}{79}$ et $\frac{\sigma^2}{50}$ par $\frac{\sigma^2_B}{49}$.

L'écart réduit de la variable aléatoire d est égal à :

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\sigma^2_A}{79} + \frac{\sigma^2_B}{49}}} = \frac{80 - 78}{\sqrt{\frac{100}{79} + \frac{81}{49}}} = 0{,}43.$$

La probabilité de trouver un écart réduit supérieur à 0,43 ou inférieur à — 0,43 est donnée par la table de Gauss, soit : $2(1 - 0{,}666) = 0{,}668$. Cette probabilité est beaucoup trop grande pour que la différence de rendement soit significative. Les engrais A et B ont un même rendement.



Ces quelques exemples donnent une idée du principe de la méthode des sondages. Ces méthodes ont été développées dans chaque branche de la connaissance où elles trouvent une application. De nombreuses tables et abaques permettent de lire directement les résultats.
Il est impossible, dans le cadre de cet article, d'envisager toutes les possibilités de la statistique mathématique.
Nous avons essayé de dégager quelques principes fondamentaux qui peuvent donner un aperçu du mode de raisonnement ainsi que des résultats que l'on peut attendre des statistiques. Ces résultats ne sont jamais des certitudes, mais seulement des indications très précieuses pour le chercheur ou le technicien.

Michèle Pascal.

STRICT, STRICTEMENT ◆ page 259.

◆ Voir Ordre.

Un nombre réel a est strictement inférieur (ou supérieur) à un nombre réel b si a est inférieur (ou supérieur) à b, l'égalité entre a et b étant exclue.
Ainsi, un nombre réel x est strictement positif s'il est positif sans être nul ($x > 0$, $x \neq 0$). L'ensemble des réels strictement positifs est noté R_+^*.

STRUCTURE : voir l'article « l'Algèbre », pages 16, 27 à 54, et la page 173.

SUITE (arithmétique, géométrique, harmonique) : voir l'article « l'Analyse », pages 107 à 114, et les pages 338, 408, 429, 452.

SUITE DE CAUCHY ◆

◆ Lire d'abord « les Suites » dans l'article « l'Analyse », pages 107 et suivantes.

Considérons la suite de nombres réels, définie, pour tout entier positif n, par :

$$u_n = \frac{1}{n}.$$

Si p et q sont deux entiers positifs (tels que $p < q$), la différence entre $\frac{1}{p}$ et $\frac{1}{q}$ ◆ peut être rendue aussi proche de 0 que l'on veut, à condition de choisir p et q suffisamment grands. Par exemple, elle peut être rendue inférieure à $\frac{1}{1\,000} = 0{,}001$ si on choisit p et q supérieurs à 1 000. Elle peut être rendue inférieure à $\frac{1}{100\,000} = 0{,}000\,01$ si on choisit p et q supérieurs à 100 000.

◆ Cette différence $\frac{1}{p} - \frac{1}{q}$ entre le plus grand des deux nombres et le plus petit est la distance de ces deux nombres.

On dit que la suite $u_n = \frac{1}{n}$ est une suite de Cauchy.

De façon générale, si E est un espace muni d'une distance d, une suite x_n de points de E est une suite de Cauchy si, quel que soit le nombre réel ε strictement positif, il est possible de trouver un entier N tel que, si p et q sont deux entiers quelconques supérieurs à N, alors la distance de x_p à x_q est inférieure à ε :

$$\forall\, \varepsilon > 0,\ \exists\, \mathrm{N} \in \mathbb{N} \mid p > \mathrm{N} \text{ et } q > \mathrm{N} \Longrightarrow d(x_p, x_q) < \varepsilon.$$

Propriétés

Toute suite convergente est une suite de Cauchy, mais la réciproque n'est pas vraie dans n'importe quel espace métrique ◆. Un espace métrique pour lequel toute suite de Cauchy converge est dit complet. Dans $\mathbb{R}$, toute suite de Cauchy converge (la suite vue plus haut, $u_n = \frac{1}{n}$, converge dans $\mathbb{R}$; sa limite est 0). Par contre, si on considère l'ensemble E qui est égal à l'intervalle]0, 1] (muni de la même distance que $\mathbb{R}$), la suite $u_n = \frac{1}{n}$ est une suite de Cauchy de points de E, mais elle ne converge pas dans E (puisque sa limite serait 0 et que 0 n'appartient pas à E).

◆ *espace métrique* : ensemble sur lequel est définie une distance.

Remarque

Si un ensemble peut être muni d'une distance de plusieurs façons différentes (voir le cas de $\mathbb{R}^2$ qui peut l'être d'au moins trois façons), il est possible que, pour l'une des distances, il soit complet, mais pas pour une autre.

SUPÉRIEUR : voir Inférieur et Ordre

SUPPLÉMENTAIRE

Si E est un espace vectoriel, deux sous-espaces vectoriels E_1 et E_2 de E sont supplémentaires si :
— leur intersection ne comporte que le vecteur nul : $E_1 \cap E_2 = \{\vec{0}\}$;
— tout vecteur de E peut s'écrire sous forme de la somme d'un vecteur de E_1 et d'un vecteur de E_2 (ces deux vecteurs sont d'ailleurs déterminés de manière unique) : $\forall \vec{x} \in E, \ \exists \vec{x_1} \in E_1, \ \exists \vec{x_2} \in E_2, \ \vec{x} = \vec{x_1} + \vec{x_2}$ ◆.

◆ Lire : quelque soit le vecteur $\vec{x}$ de E, il existe un vecteur $\vec{x_1}$ de E_1 et un vecteur $\vec{x_2}$ de E_2 tels que $\vec{x}$ soit égal à la somme $\vec{x_1} + \vec{x_2}$.

Par exemple, l'ensemble $\mathbb{R}^2$ des couples (x, y) de nombres réels est un espace vectoriel sur $\mathbb{R}$. Les deux sous-ensembles E_1 et E_2, définis de la façon suivante :

E_1 est l'ensemble des couples $(x, 0)$
E_2 est l'ensemble des couples $(0, y)$,

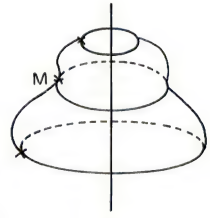
sont deux sous-espaces vectoriels de $\mathbb{R}^2$ (cela se démontre facilement). Leur intersection ne comporte que le couple (0, 0), et tout couple (x, y) peut se mettre sous la forme de la somme d'un couple de E_1 et d'un couple de E_2 :

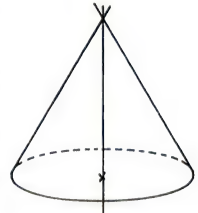
$$(x, y) = (x, 0) + (0, y).$$

Ces deux sous-espaces E_1 et E_2 sont supplémentaires.

SURFACE DE RÉVOLUTION

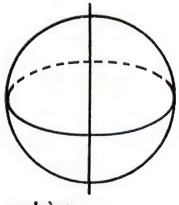
Surface engendrée par une courbe qui tourne autour d'une droite. La courbe est appelée génératrice et la droite, axe de révolution de la surface.
Chaque point de la génératrice parcourt un cercle, dans un plan perpendiculaire à l'axe.
Parmi les surfaces de révolution, on peut citer :
— *le cône de révolution*, pour lequel une courbe génératrice est une droite qui coupe l'axe en un point qui est le sommet du cône ;
— *le cylindre de révolution*, pour lequel une courbe génératrice est une droite parallèle à l'axe ;
— *la sphère*, qui peut être considérée comme engendrée par la rotation d'un demi-cercle autour de son diamètre ;
— *le tore*, dont la courbe génératrice est un cercle situé dans un plan contenant l'axe ;
— *l'ellipsoïde de révolution*, dont la courbe génératrice est une ellipse et l'axe un des axes de l'ellipse ;
— *l'hyperboloïde de révolution* à une nappe, dont une courbe génératrice est une droite n'appartenant pas au même plan que l'axe, et une autre courbe génératrice est une hyperbole dont l'axe non transverse est l'axe de révolution.



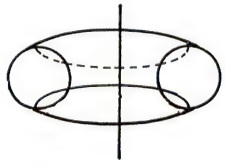


cône de révolution

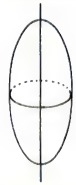
cylindre de révolution



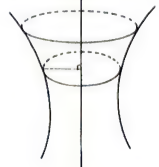
sphère



tore

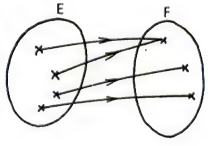


ellipsoïde de révolution



hyperboloïde de révolution

SURJECTION pages 80, 206, 277 et 278, 335.



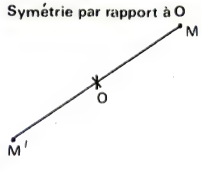
Une application f d'un ensemble E vers un ensemble F est une surjection de E sur F (ou est surjective) si tout élément de F est image par f d'au moins un élément de E : $\forall y \in F, \ \exists x \in E, \ y = f(x)$

(lire : quel que soit y appartenant à l'ensemble F, il existe au moins un élément x de E tel que $y = f(x)$).

Exemples

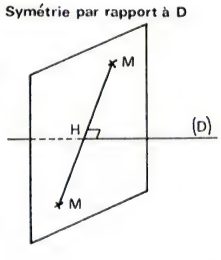
L'application de $\mathbb{R}$ dans $\mathbb{R}$ (ensemble des nombres réels) qui à x fait correspondre $2x$ est surjective, celle de $\mathbb{N}$ dans $\mathbb{N}$ (ensemble des entiers naturels positifs) qui à x fait correspondre $x + 1$ n'est pas surjective, car 0, élément de $\mathbb{N}$, n'est image d'aucun entier (il n'y a pas de nombre entier positif tel qu'en lui ajoutant 1 on obtienne 0).

SYMÉTRIE pages 243, 281, 298, 300.



Symétrie par rapport à un point O

C'est la transformation géométrique qui à tout point M de l'espace fait correspondre le point M′ tel que O soit le milieu de MM′. On dit que M et M′ sont symétriques par rapport à O.



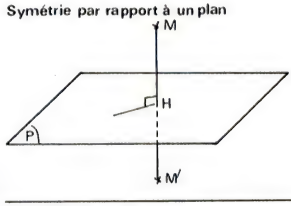
Symétrie par rapport à une droite D

C'est la transformation géométrique qui à chaque point M de l'espace associe le point M′ symétrique de M par rapport au point H (intersection de la droite D avec le plan perpendiculaire à D passant par M).

Symétrie par rapport à un plan P

C'est la transformation géométrique qui à chaque point M associe le point M′ symétrique de M par rapport au point H (intersection de P avec la droite perpendiculaire à P passant par M).

Ces trois symétries possèdent la propriété suivante : la composée de l'une quelconque d'entre elles avec elle-même est la transformation identique (car le symétrique du symétrique de M est M).



SYMÉTRIQUE (Elément) : voir l'article « l'Algèbre », page 23.

SYMÉTRIQUE (Matrice)

Une matrice carrée est symétrique si deux éléments quelconques occupant deux positions symétriques par rapport à la diagonale principale de la matrice sont égaux ◆.

$$\begin{pmatrix} 1 & -1 & 3 \\ -1 & 0 & -2 \\ 3 & -2 & 2 \end{pmatrix}$$

Matrice symétrique

SYSTÈME D'ÉQUATIONS

Un système d'équation est un ensemble de plusieurs équations. Résoudre un système, c'est trouver toutes les valeurs des inconnues pour lesquelles les équations sont simultanément vérifiées.

Système de deux équations linéaires à deux inconnues : x et y

C'est un système de la forme :

$$\begin{cases} ax + by = c \\ a'x + b'y = c' \end{cases}$$

où a, b, c, a', b', c' sont des constantes et où les inconnues sont x et y. Dans le cas où $c = c' = 0$, le système est dit homogène.

— Si le déterminant $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - ba'$ du système est différent de 0, il y a une seule solution : le couple (x, y) tel que :

$$\begin{cases} x = \dfrac{cb' - c'b}{ab' - ba'} = \dfrac{\begin{vmatrix} c & b \\ c' & b' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}} \\ y = \dfrac{ac' - ca'}{ab' - ba'} = \dfrac{\begin{vmatrix} a & c \\ a' & c' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}} \end{cases}$$

Le système est dit système de Cramer.

Remarquons que, dans le cas d'un système homogène à déterminant non nul, il y a une seule solution : $x = 0$ et $y = 0$.

— Si $ab' - ba'$ est nul et :

- si $ac' - a'c$ n'est pas nul, alors il n'y a pas de solution ;
- si $ac' - a'c$ est nul (alors $cb' - c'b$ l'est aussi), il y a une infinité de solutions : tous les couples (x, y) vérifiant $ax + by = c$ (la seconde équation est alors la même que la première).

Exemples

1) $\begin{cases} 2x + 3y = 5 \\ 3x + 5y = 2 \end{cases}$ $\quad \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix} = 2 \times 5 - 3 \times 3 = 1.$

Il y a une seule solution :

$$x = \frac{\begin{vmatrix} 5 & 3 \\ 2 & 5 \end{vmatrix}}{1} = 25 - 6 = 19 \qquad y = \frac{\begin{vmatrix} 2 & 5 \\ 3 & 2 \end{vmatrix}}{1} = 4 - 15 = -11.$$

2) $\begin{cases} 2x + 3y = 5 \\ -4x - 6y = 2 \end{cases}$ $\quad \begin{vmatrix} 2 & 3 \\ -4 & -6 \end{vmatrix} = -12 + 12 = 0,$

mais $\begin{vmatrix} 2 & 5 \\ -4 & 2 \end{vmatrix} = 4 + 20 = 24 \neq 0.$

Il n'y a pas de solution (le premier membre de la deuxième équation est égal au premier membre de la première multiplié par — 2, mais le second membre de la deuxième équation n'est pas égal au second membre de la première multiplié par — 2).

3) $\begin{cases} 2x + 3y = 5 \\ -4x - 6y = -10 \end{cases}$ $\quad \begin{vmatrix} \begin{vmatrix} 2 & 3 \\ -4 & -6 \end{vmatrix} = 0 \\ \begin{vmatrix} 2 & 5 \\ -4 & -10 \end{vmatrix} = 0 \\ \begin{vmatrix} 3 & 5 \\ -6 & -10 \end{vmatrix} = 0. \end{vmatrix}$

Les deux équations sont, en fait, proportionnelles (la seconde est égale à la première multipliée par — 2) ; il y a une infinité de solutions : tous les couples (x, y) vérifiant l'une ou l'autre de ces équations.

On peut généraliser ces résultats à un système de n équations linéaires à p inconnues.

T

TABLE DE PYTHAGORE : voir l'article « l'Algèbre », et les pages 18 à 21, 29, 36.

TABLE DE VARIATION : voir l'article « l'Analyse », pages 120, 121 et la page 312.

TABLE DE VÉRITÉ : pages 299, 302, 310, 337, 365, 396.

En logique mathématique « classique », chaque proposition ne peut être que vraie ou fausse ; il n'y a pas d'autre choix possible. Ces deux valeurs : vraie (représentée par V ou 1) et fausse (représentée par F ou 0) sont appelées les valeurs de vérité de la proposition. Si l'on met en jeu une proposition, il n'y a que deux possibilités : V ou F ; si l'on met en jeu deux propositions : p et q, il y a quatre possibilités p est vraie et q aussi, p est vraie et q fausse, p est fausse et q est vraie, p est fausse et q aussi. On peut disposer ces quatre possibilités dans un tableau représenté ci-contre.

p	q
V	V
V	F
F	V
F	F

Si on veut faire un calcul quelconque sur ces deux propositions, on peut compléter le tableau par une troisième colonne, dans laquelle on mettra la valeur de la proposition qui fait intervenir p et q dans chacun des quatre cas possibles. Un tel tableau est appelé « table de vérité ». Voir ci-contre la nouvelle table de vérité de la proposition « p ou q » (proposition vraie si l'une au moins des deux propositions p, q est vraie).

p	q	p ou q
V	V	V
V	F	V
F	V	V
F	F	F

Reprenons le cas d'une seule proposition p à partir de laquelle on définit une autre proposition (par exemple, la négation de p ◆ qui est vraie lorsque p est fausse et fausse lorsque p est vraie), la table de vérité ne comportera que deux lignes.
Si l'on met en jeu trois propositions p, q, r, il y a huit possibilités résumées dans le tableau ci-dessous :

◆ Notée non p ou $\neg p$.

p	non p
V	F
F	V

p	q	r	
V	V	V	les trois propositions sont vraies
V	V	F	deux propositions sont vraies et une est fausse
V	F	V	
F	V	V	
V	F	F	une proposition est vraie et deux sont fausses
F	V	F	
F	F	V	
F	F	F	les trois propositions sont fausses

Si on veut faire un calcul à l'aide de ces trois propositions, la table de vérité comportera donc huit lignes.
On peut utiliser ces tables de vérité pour faire, de façon simple et rapide, certaines démonstrations. Par exemple, supposons vouloir démontrer que la négation de la proposition « p et q » ◆ est équivalente ◆ à la proposition « (non p) ou (non q) ◆ ».
On fait une table de vérité dans laquelle on va mettre sept colonnes : p, q, p et q, non (p et q), non p, non q, non p ou non q ; la comparaison de la quatrième et de la septième colonne permettra de conclure à l'équivalence ou non de ces deux propositions.

◆ La proposition « p et q » est celle qui est vraie si et seulement si p et q sont vraies en même temps.

◆ Deux propositions sont équivalentes si elles sont vraies en même temps et fausses en même temps.

◆ La proposition « p ou q » est la proposition qui est vraie si et seulement si l'une au moins des deux propositions p ou q est vraie.

p	q	p et q	non (p et q)	non p	non q	non p ou non q
V	V	V	F	F	F	F
V	F	F	V	F	V	V
F	V	F	V	V	F	V
F	F	F	V	V	V	V

Les deux propositions non (p et q) et (non p) ou (non q) sont équivalentes puisque vraies en même temps et fausses en même temps.

TANGENT (Plan)

Soient une surface (S) et un point A de cette surface. Si toutes les courbes tracées sur la surface (S) et passant par A admettent une tangente en A, en général toutes ces tangentes appartiennent à un même plan appelé plan tangent en A à la surface (S).
Une sphère admet en chacun de ses points un plan tangent qui est le plan passant par ce point et perpendiculaire au rayon qui aboutit en ce point.
Si l'équation de la surface est $f(x, y, z) = 0$, l'équation du plan tangent en A de coordonnées (x_0, y_0, z_0) est :

$$(x - x_0)\frac{\partial f}{\partial x}(x_0, y_0, z_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0, z_0) + (z - z_0)\frac{\partial f}{\partial z}(x_0, y_0, z_0) = 0$$

$\frac{\partial f}{\partial x}(x_0, y_0, z_0)$, $\frac{\partial f}{\partial y}(x_0, y_0, z_0)$ et $\frac{\partial f}{\partial z}(x_0, y_0, z_0)$ sont les valeurs des dérivées partielles de f obtenues en y remplaçant x, y, z par les coordonnées de A.

TANGENTE page 327.

Tangente à une courbe

La tangente en A à la courbe (C) est la limite, si elle existe, de la droite AM (joignant A à un point M quelconque de la courbe) quand M tend vers A.

Exemple : la tangente à un cercle en un point A de ce cercle est la droite perpendiculaire en A au rayon OA.

Recherche de l'équation d'une tangente

1) *Si la courbe est définie par son équation cartésienne* $y = f(x)$ et si x_0 et y_0 sont les coordonnées du point A, l'équation de la tangente en A à (C) est :

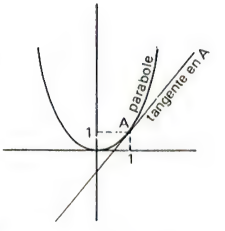
$$y - y_0 = f'(x_0)(x - x_0)$$

où $f'(x_0)$ est la valeur en x_0 de la dérivée de f.

Exemple : la parabole d'équation $y = x^2$ admet au point A de coordonnées (1, 1) une tangente qui est la droite d'équation :

$$y - 1 = 2(x - 1)$$

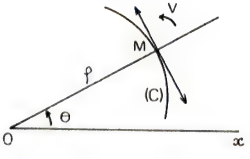
car la dérivée de x^2 est $2x$: pour $x = 1$, $2x$ vaut 2.



2) *Si la courbe est définie par des équations paramétriques* $x = f(t)$, $y = g(t)$. f et g étant deux fonctions dérivables, la tangente au point A (correspondant à la valeur t_0 de t) a pour équation :

$$\frac{x - x_0}{f'(t_0)} = \frac{y - y_0}{g'(t_0)}$$

où $f'(t_0)$ et $g'(t_0)$ sont les valeurs des dérivées de f et g pour t_0.



3) *Si la courbe est définie par son équation polaire* $\rho = f(\theta)$, l'angle appelé V formé par OM et la tangente en M à la courbe est tel que :

$$\text{tg } V = \frac{\rho}{\rho'} \text{ (où } \rho' \text{ est la dérivée de } \rho).$$

L'équation polaire de la tangente en M à la courbe est :

$$\frac{1}{\rho} = \frac{1}{\rho_0}\cos(\theta - \theta_0) + \left(\frac{1}{\rho}\right)'_0 \sin(\theta - \theta_0)$$

où ρ_0 et θ_0 sont les coordonnées polaires de M et $\left(\frac{1}{\rho}\right)'_0$ est la valeur de la dérivée de $\frac{1}{\rho}$ obtenue en remplaçant θ par θ_0.

Exemple de calcul : tangente en $M(\rho_0, \theta_0)$ au cercle de centre A passant par O (R étant son rayon) ; l'équation de ce cercle est :

$$\rho = 2R\cos\theta$$

$$\left(\frac{1}{\rho}\right)' = \left(\frac{1}{2R\cos\theta}\right)' = +\frac{\sin\theta}{2R\cos^2\theta}.$$

Pour θ_0, cette dérivée vaut : $\dfrac{\sin\theta_0}{2R\cos^2\theta_0}$.

L'équation de la tangente est :

$$\frac{1}{\rho} = \frac{1}{2R\cos\theta_0}\cos(\theta - \theta_0) + \frac{\sin\theta_0}{2R\cos^2\theta_0}\sin(\theta - \theta_0) = \frac{\cos(2\theta_0 - \theta)}{2R\cos^2\theta_0}.$$

Cas d'une courbe gauche ◆

◆ Courbe non plane, dans l'espace.

La définition est la même. Si la courbe est définie par ses équations paramétriques $x = f(t)$, $y = g(t)$, $z = h(t)$, les équations de la tangente en A (correspondant à la valeur t_0 du paramètre) sont :

$$\frac{x - x_0}{f'(t_0)} = \frac{y - y_0}{g'(t_0)} = \frac{z - z_0}{h'(t_0)}.$$

Fonction tangente : voir Trigonométriques (fonctions).

Courbes tangentes
Deux courbes (C) et (C′) sont dites tangentes en M si (C) et (C′) ont la même tangente au point M.

Surfaces tangentes
Deux surfaces (S) et (S′) sont tangentes en M si elles ont le même plan tangent en M.

TÉTRAÈDRE : voir Polyèdres et la page 320.

TOPOLOGIE pages 175, 186, 307, 330.

La topologie est sans doute l'un des domaines de la mathématique qu'il nous est le plus difficile de caractériser de façon qui soit à la fois concise et précise. Difficulté qui a sans doute plusieurs explications, ne serait-ce que l'aspect innovant de cette discipline et ce qu'elle peut avoir, par conséquent, de déroutant. Cette sorte de résistance inévitable de la pensée à s'assimiler un nouveau type de réflexion et de recherche se renforce ici, dans nos esprits, par l'habitude — suscitée par tout un passé scolaire — qui consiste à concevoir le tronc mathématique comme ramifié une fois pour toutes en deux grandes branches : l'algèbre et la géométrie. De ce fait, rien de plus déconcertant que d'ouvrir un livre de topologie. A simplement le feuilleter, on a la surprise d'y découvrir aussi bien des figures plus ou moins familières dont on envisage les déformations possibles que des pages entières d'écriture symbolique qui, au regard de la tradition, évoqueraient plutôt l'algèbre : en effet, le discours qu'elle tient ressortit aussi bien de l'étude des fonctions que de la géométrie des surfaces. Si ces deux champs, jadis séparés (quoique partiellement rapprochés par la géométrie analytique), peuvent être aujourd'hui réunis dans une même approche, une même perspective, cela ne peut être qu'à la faveur d'une abstraction commune qui, en fait, les unifie. Cette abstraction, c'est la topologie. Ce qu'elle crée de commun et d'abstrait, c'est un langage. Langage formalisé tel qu'il devient équivalent de parler de fonctions ou des propriétés « géométriques » de l'ensemble des nombres réels, par exemple. Il serait donc erroné, en tout cas réducteur, de s'en tenir à l'étymologie du mot (venant du grec : « topos ») pour en faire un discours sur le lieu, la science des lieux, en un sens qui ferait de la topologie un simple synonyme de la géométrie la plus traditionnelle.

Historique

Sans doute les premières manifestations de la pensée auxquelles on puisse épingler le qualificatif de topologiques sont-elles largement apparentées à la géométrie. Ainsi, si l'on se réfère aux problèmes qui ont préoccupé Leonhard Euler ◆ (1707-1783), le grand-père de la topologie, on trouve le problème classique des sept ponts de Kœnigsberg, dont on disait qu'ils ne pouvaient être successivement traversés sans que le promeneur soit obligé de repasser par l'un des ponts déjà empruntés. Euler donna une démonstration mathématique de cette impossibilité en des termes qui préfiguraient la notion topologique moderne de chemin. Il établit également que, si F est le nombre des faces d'un polyèdre, S le nombre des sommets et A le nombre d'arêtes, on a toujours $F + S - A = 2$. A ce stade, la topologie est donc encore assimilable à la géométrie. Elle va apparaître

◆ L. Euler (1707-1783) : le plus grand homme de science suisse, mathématicien très prolixe.

en tant que telle à partir d'une certaine réflexion portant sur les fonctions, c'est-à-dire qu'elle va émerger du champ de l'analyse ◆. Cette émergence, on peut la repérer dans l'œuvre de Karl Weierstrass (1815-1897), qui, vers 1860, découvre et étudie la notion de limite d'une fonction. Cela le conduit à approfondir, à systématiser la structure de l'ensemble de nombres sur lequel portent les fonctions numériques, à savoir les nombres réels et, par là même, il est amené à mettre au point des propriétés de cet ensemble qui entrent pleinement dans le cadre de ce que l'on appellerait aujourd'hui ses propriétés topologiques. A commencer par la construction qu'il donne de l'ensemble $\mathbb{R}$ des nombres réels par « coupures » dans l'ensemble des nombres rationnels.

◆ Ce terme regroupe classiquement l'étude des fonctions et l'ensemble du calcul infinitésimal.

Le second pas important franchi par la topologie est dû à Georg Cantor (1845-1918). Fort de sa découverte de la théorie des ensembles (qui structure aujourd'hui toute l'architecture de notre mathématique), Cantor étudie en particulier les ensembles de points. Il se situe ainsi comme le pionnier d'un des grands courants de la topologie que l'on désigne du nom de « topologie ensembliste ». C'est à Henri Poincaré qu'il revient d'en avoir fixé les bases à la fin du siècle dernier : il s'agit de la « topologie algébrique » ou « combinatoire », dont l'objectif essentiel est la recherche des « invariants topologiques » d'une configuration (voir plus loin).

Indiquons encore que la topologie ensembliste fut élaborée et systématisée par F. Hausdorff vers 1900-1910, qu'une synthèse des topologies ensembliste et combinatoire fut réalisée peu après L.E.J. Brower. Mais les résultats cohérents n'apparaîtront qu'ultérieurement vers 1930 et seront dus à J.W. Alexander, P.L. Alexandrov et S. Lefschetz. C'est d'ailleurs à ce dernier qu'il revient d'avoir introduit le vocable de « topologie ». Jusque-là, on ne parlait que d'« analyse locale », ce qui indiquait bien la marque d'origine de la topologie. Depuis 1930, la topologie n'a cessé de croître et d'embellir, revivifiant des secteurs importants de la mathématique. Elle a également pénétré dans d'autres secteurs scientifiques que la mathématique et joue notamment un rôle décisif dans la formalisation des conceptions relativistes.

Aujourd'hui, la topologie est une théorie autonome qui, certes, n'a pas le même coefficient de généralisation que la théorie des ensembles, mais qui, comme elle, est susceptible de multiples applications.

L'objet de la topologie : la continuité

Ce qu'il faut retenir, c'est que la topologie provient en ligne directe des besoins de l'analyse qui l'ont fait naître et qu'à son tour elle est à même de rendre compte d'un certain nombre de problèmes d'analyse, c'est-à-dire portant sur les fonctions. A cet égard, elle est en quelque sorte utilisable comme un modèle abstrait, dans la mesure où elle permet une simulation géométrique d'un certain nombre de questions qui n'étaient traitables auparavant que dans le seul champ de l'algèbre des fonctions et du calcul infinitésimal. Cette idée d'un modèle, susceptible de rendre compte de multiples réalités mathématiques, explique ce que Kuratowski écrit de la topologie : « Encore récemment une agglomération de théorèmes » mal reliés entre eux, elle est devenue un système mathématique et les méthodes » topologiques ont pénétré dans beaucoup d'autres domaines des mathéma- » tiques. » En guise d'exemple, il ajoute qu'ainsi des théorèmes sur l'existence d'une solution pour certains types d'équations différentielles peuvent être exprimés comme des théorèmes sur l'existence de points d'un espace de fonctions (les fonctions continues) invariantes pour certaines transformations continues.

Il va de soi que cette simulation topologique de problèmes analytiques n'a de

sens que si l'on réalise du même coup une simplification de la question à traiter. Or tel est justement le cas : la topologie opère avec des concepts plus généraux que ceux de l'analyse. De ce fait, elle fournit souvent une approche adéquate pour des problèmes auxquels l'analyse seule ne peut donner réponse.

La fécondité de la démarche, la généralité des méthodes topologiques tiennent à la généralité des ensembles auxquels elles peuvent être appliquées. En effet, il sera montré par la suite qu'il est possible d'introduire une structure topologique non pas seulement dans des ensembles géométriques classiques, c'est-à-dire des ensembles de points, mais qu'il est possible également de l'appliquer à des objets aussi peu « géométriques », en première approximation, que fonctions ou suites numériques.

De façon schématique, le modèle topologique est susceptible de rendre compte essentiellement de deux grands champs : certaines propriétés des ensembles de points, d'une part, et les propriétés (de même nature) relatives aux ensembles de fonctions, d'autre part.

Ces deux champs peuvent sembler au départ profondément hétérogènes. Une notion essentielle pourtant les rapproche. Et à préciser cette notion, nous tiendrons du même coup l'objet même de la topologie, au sens où l'on parle de l'objet d'une science comme étant ce qui la fonde : l'objet de la topologie, c'est la continuité.

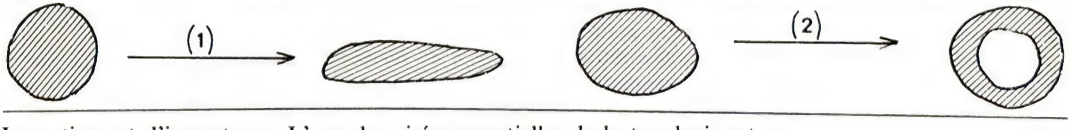
Il est alors possible de comprendre en quoi les deux champs que nous avons distingués peuvent se chevaucher, se rejoindre dans le cadre de la topologie, La continuité est en effet une propriété décisive des fonctions numériques.

Une géométrie élastique

Quant au champ plus proprement géométrique, la continuité y apparaîtra essentiellement comme caractéristique de ce qu'on appelle justement les transformations continues, lesquelles offrent un intérêt privilégié pour le topologue.

Il nous faut d'abord préciser ce qu'on entend par transformation continue d'une configuration géométrique. On peut en donner une illustration de physique intuitive en imaginant les manipulations que l'on peut faire subir à une masse de pâte à modeler sans y introduire de ruptures ou de cassures. Les transformations continues sont donc les transformations qui altèrent, modifient la surface sans y faire apparaître de discontinuité.

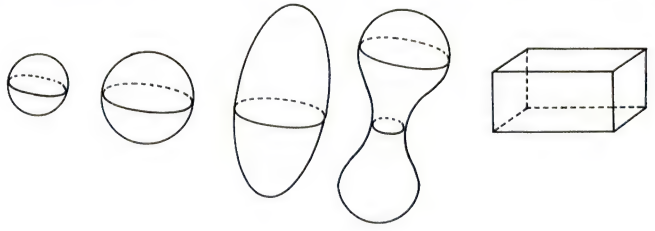
Par exemple, le passage (1) de la boule au bâtonnet est une transformation continue, le passage (2) de la boule à l'anneau n'en est pas une, car on a fait apparaître un trou.



La notion est d'importance. L'une des visées essentielles de la topologie est en effet l'« étude des propriétés des configurations géométriques qui restent inva- » riantes lorsque ces configurations sont soumises à des transformations biunivoques » et bicontinues », c'est-à-dire continues quand on passe de l'état (*a*) à l'état (*b*) et, réciproquement, de l'état (*b*) à l'état (*a*).

C'est cette orientation qui a permis de voir dans la topologie une « géométrie élastique ». Du point de vue topologique sont considérés comme équivalents tous les ensembles de points tels qu'on puisse passer de l'un à l'autre par déformation continue.

Pour le topologue, les figures suivantes sont équivalentes :



Il devient alors possible de fixer pour la suite quelques-unes des grandes notions qui structurent le champ de la topologie.
Nous avons déjà indiqué la première : celle de transformation continue, qu'on appelle encore homéomorphisme. Deux figures sont dites topologiquement équivalentes (au sens de la relation d'équivalence définie en théorie des ensembles) ou homéomorphes si chacune d'elles peut se déduire de l'autre par transformation continue ou homéomorphisme.

De ce fait, la topologie s'intéresse essentiellement aux propriétés intrinsèques des figures, c'est-à-dire aux propriétés qui ne font pas référence à l'espace environnant dans lequel les figures en question peuvent se trouver placées ; plus précisément : aux propriétés qui ne sont pas altérées par un homéomorphisme. De telles propriétés sont appelées invariants topologiques ◆.

◆ Ainsi, un cercle et un triangle sont deux figures homéomorphes. La propriété qui consiste à partager le plan en deux régions (l'intérieur et l'extérieur de la figure) vaut pour l'une comme pour l'autre. C'est un invariant topologique. Par contre, la propriété qu'a le cercle d'admettre une tangente en tout point ne se retrouve pas dans le triangle (pour les sommets). Elle n'est donc pas un invariant.

L'espace topologique

Mais en rester là, à l'examen de figures classiques (cercle, triangle...), voire moins habituelles (bande de Möbius, bouteille de Klein), c'est demeurer au niveau « naïf » de la topologie. Or, nous l'avons dit, la fécondité de la méthode tient à la possibilité d'introduire des considérations topologiques dans un ensemble quel qu'il soit. Tout ensemble peut être « topologisé ». C'est même par ce biais, décisif, que la topologie peut être raccrochée à la théorie des ensembles, voire fondée à partir de celle-ci.
Dans ce processus, qui est processus d'abstraction, le pas décisif que franchit la topologie par rapport à la géométrie classique consiste à s'affranchir de la notion de distance. En effet, du point de vue topologique, la notion de distance est inadéquate, puisqu'elle n'est pas un invariant : la distance de deux points d'une figure pouvant varier lors d'une déformation continue. Il faut donc fonder la notion d'espace sur une base moins rigide, plus abstraite. C'est ce à quoi pourvoit la notion d'espace topologique (définie ultérieurement).
Dans la perspective ouverte par ce nouveau concept, l'espace euclidien classique à trois dimensions, muni de la distance au sens habituel du terme, est retrouvé comme un simple cas particulier. Mais on s'en est considérablement éloigné au sens d'une abstraction généralisatrice.
On l'a compris : c'est justement cette généralisation qui fait de la topologie aussi bien un moyen d'investigation des courbes et des surfaces, qu'elles appartiennent aux espaces classiques à une, deux ou trois dimensions ou à un espace à n dimensions, que de l'ensemble des nombres réels sur lesquels portent, opèrent les fonctions de l'algèbre classique.

L'espace topologique de l'ensemble des nombres réels

Au nombre des fonctions, celles qui relèveront au premier chef de la topologie seront les fonctions continues dans la mesure où elles réalisent une déformation continue de l'ensemble $\mathbb{R}$ des nombres réels. La connaissance des propriétés topologiques de $\mathbb{R}$ qu'éclairent les notions de continuité, naturellement, mais aussi de limites, de points limites, de points d'accumulation, de voisinages, d'ouverts, etc., sont riches d'implications dans le domaine du calcul infinitésimal. A mieux comprendre l'ensemble sur lequel elles opèrent, on comprend davantage les propriétés infinitésimales des fonctions ; et ce, indépendamment de la notion classique de distance, à laquelle s'est substituée celle d'« ouvert », qui suffit à elle seule à rendre compte de la continuité.

De même, un théorème d'algèbre comme le suivant : « Un polynôme de degré impair à coefficients réels admet au moins un zéro réel ◆ », entre aisément dans le cadre de considérations topologiques (il s'explique aisément par la continuité). C'est bien ce qui rend si séduisante, si enrichissante la démarche topologique. Non seulement elle permet d'expliquer des propriétés d'algèbre et de géométrie classiques, mais, outre qu'elle les prolonge largement, elle permet l'étude de secteurs mathématiques bien plus abstraits qui caractérisent le vertigineux éventail des découvertes de la mathématique contemporaine.

Gérard Guillerault.

◆ C'est-à-dire qu'un polynôme du type $ax^3 + bx^2 + cx + d$, par exemple, où a, b, c, d sont des nombres réels, s'annule nécessairement pour une valeur réelle de x. Autrement dit, il existe un nombre réel x tel que $ax^3 + bx^2 + cx + d = 0$.

Définition mathématique d'une topologie sur un ensemble ◆

Soit E l'intervalle semi-ouvert $[0, 1[$ de l'ensemble des nombres réels $\mathbb{R}$. Soit F la famille constituée des sous-ensembles suivants de E :

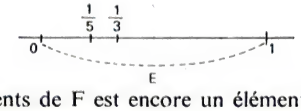
— l'ensemble vide $\varnothing$;

— les intervalles $[0, 1[$, $\left[0, \frac{1}{2}\right[$, $\left[0, \frac{1}{3}\right[$, ..., c'est-à-dire tous les intervalles de la forme $\left[0, \frac{1}{n}\right[$, où n est un entier positif non nul ($n \in \mathbb{N}^*$).

1) L'ensemble vide et l'ensemble E lui-même appartiennent à cette famille F.

2) La réunion d'un nombre quelconque d'éléments de F fait encore partie de F. Par exemple :

$$\varnothing \cup \left[0, \frac{1}{3}\right[\cup \left[0, \frac{1}{5}\right[= \left[0, \frac{1}{3}\right[.$$



3) L'intersection d'un nombre fini d'éléments de F est encore un élément de F. Par exemple :

$$\left[0, \frac{1}{4}\right[\cap \left[0, \frac{1}{5}\right[\cap \left[0, \frac{1}{6}\right[\cap \ldots \cap \left[0, \frac{1}{n}\right[= \left[0, \frac{1}{n}\right[.$$

◆ Pour le lecteur peu familiarisé avec les notions d'ensemble, il est conseillé de lire d'abord l'article « les Ensembles » et Intervalle. Pour compléter la notion de topologie, lire aussi Accumulation, Adhérence, Adjacent, Base, Bon ordre, Boule, Compact, Complet, Connexe, Dense, Distance, Fermé, Filtre, Frontière, Homéomorphisme, Induite, Intérieur, Isolé, Isométrie, Normé, Orthogonal, Orthonormé, Recouvrement, Séparé, Voisinage.

Ouvert : espace topologique

Si on appelle *ouverts* de E tous les intervalles de F, on dit qu'on a défini une topologie sur E ou encore que E est un *espace topologique*.

De manière générale, on dit qu'un ensemble E est un espace topologique si on a défini une famille F de parties de E telles que :

— Propriété 1 : $\varnothing$ et E appartiennent à F.

— Propriété 2 : toute réunion d'éléments de F est élément de F.

— Propriété 3 : toute intersection d'un nombre fini d'éléments de F est un élément de F.
Les éléments de F sont appelés les ouverts de E.

Topologies grossière et discrète

On peut toujours munir un ensemble quelconque E de l'une des deux topologies suivantes :

— topologie grossière : topologie pour laquelle les seuls ouverts de E sont la partie vide et E tout entier ;

— topologie discrète : topologie pour laquelle toute partie de E est un ouvert de E.

Exemples de topologie sur l'ensemble des nombres réels $\mathbb{R}$

Topologie de l'ordre

La topologie pour laquelle les ouverts sont : l'ensemble vide, l'ensemble des nombres réels $\mathbb{R}$ et toute réunion d'intervalles ouverts ◆, est appelée topologie de l'ordre ◆.
$]0, 1[\cup]3, 6[\cup]4, 5[$ est un ouvert pour cette topologie.

◆ Pour cette topologie, les intervalles ouverts de $\mathbb{R}$ sont des ouverts. Rappelons qu'un intervalle ouvert est noté $]a, b[$: c'est l'ensemble des nombres réels x compris strictement entre a et b.

Autre exemple de topologie sur $\mathbb{R}$

Celle dont les ouverts sont :

— l'ensemble vide $\varnothing$;

— l'ensemble des réels tout entier $\mathbb{R}$;

— tout sous-ensemble de $\mathbb{R}$ constitué par les nombres réels supérieurs à un nombre réel donné ◆.

◆ On ne peut pas prendre simplement pour seuls ouverts : $\varnothing$, $\mathbb{R}$ et tous les intervalles ouverts, car la propriété 2 d'un ouvert ne serait pas satisfaite (la réunion de deux intervalles ouverts de $\mathbb{R}$ n'étant pas nécessairement un intervalle ouvert de $\mathbb{R}$) ; ainsi : $]0, 1[\cup]2, 4[$ n'est pas un intervalle ouvert de $\mathbb{R}$, ce n'est même pas un intervalle.

Le voisinage

A partir de cette définition d'une topologie, on peut définir les voisinages V des points de l'espace topologique E, mais on peut également suivre le chemin inverse et définir une topologie sur E en définissant d'abord pour chaque point x de E une famille $\mathcal{V}(x)$ de parties de E, appelées voisinages de x. Les éléments de $\mathcal{V}(x)$ doivent vérifier les axiomes suivants ◆ :

1) Toute partie de E contenant un voisinage de x est un voisinage de x.

2) Toute intersection d'un nombre fini de voisinages de x est un voisinage de x.

3) Tout voisinage de x contient x.

4) Tout voisinage V de x contient un voisinage U de x tel que V soit voisinage de chacun des points de U.

◆ Pour cette topologie, les intervalles ouverts de $\mathbb{R}$ ne sont pas des ouverts.

Alors, si on donne comme définition d'un ouvert de E : *un ouvert de* E *est une partie de* E *qui est voisinage de chacun de ses points*, à partir de ces ouverts on peut définir des voisinages de x qui ne sont autres que les éléments de la famille $\mathcal{V}(x)$.

◆ Pour les exemples, lire Voisinage.

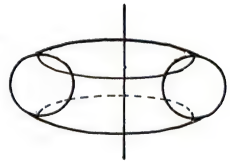
Topologie associée à une distance ◆

◆ Lire d'abord Distance et Boule.

Si E est un espace métrique, on peut munir E d'une topologie : celle dont une base d'ouverts est constituée par les boules ouvertes de E.
Un voisinage d'un point x de E est un sous-ensemble de E qui contient en même temps qu'un point une boule ouverte centrée en ce point (toute boule ouverte est un ouvert).

TORE

Surface engendrée par un cercle tournant autour d'une droite. La droite, appelée axe du tore, appartient au plan du cercle et ne recoupe pas le cercle.



TOTAL (Ordre) : voir Ordre

TRANSCENDANT (Nombre) : voir l'article « les Nombres », page 373 et 374, et la page 329.

TRANSFINI : voir Equipotent et la page 265.

TRANSFORMATION pages 251, 318, 343, 528.

Souvent, en géométrie, on appelle transformation une application du plan (ou de l'espace) dans lui-même.
Par exemple, on dit : la symétrie par rapport au point O est la transformation qui à tout point M du plan (ou de l'espace) fait correspondre le point M′ tel que O, M, M′ soient alignés et que le point O soit le milieu de MM′.

TRANSITIVE ◆ pages 45, 281, 298, 300.

◆ Lire d'abord Relation.

Si E est un ensemble dans lequel est définie une relation $\mathcal{R}$, celle-ci est transitive si, pour tout élément x de E, y de E et z de E, chaque fois que x est en relation (par $\mathcal{R}$) avec y et y en relation (par $\mathcal{R}$) avec z, alors x est en relation (par $\mathcal{R}$) avec z :

$$\forall x \in \mathrm{E}, \quad \forall y \in \mathrm{E}, \quad \forall z \in \mathrm{E} \quad (x \mathcal{R} y \text{ et } y \mathcal{R} z) \Longrightarrow x \mathcal{R} z$$

(lire : quel que soit x appartenant à E, etc.).
Par exemple, dans $\mathbb{N}$ (ensemble des entiers naturels), la relation définie par « x a même reste dans la division par 2 que y » est transitive, car, pour tout triplet (x, y, z), si x a même reste dans la division par 2 que y et y que z, alors x a même reste dans la division par 2 que z.
3 a même reste dans la division par 2 que 5 (à savoir 1) ;

5 — — — 7 —
Donc
3 — — — 7 —

Par contre, la relation « x est le double de y » (ou $x = 2y$) n'est pas transitive, car, si $x = 2y$ et $y = 2z$, alors $x = 4z$, donc x n'est pas en général le double de z.

TRANSLATION page 264.

Étant donné un vecteur $\vec{V}$, on appelle translation de vecteur $\vec{V}$ la transformation géométrique qui à tout point M du plan (ou de l'espace) fait correspondre le point M′ tel que le vecteur $\overrightarrow{MM'}$ et le vecteur $\vec{V}$ soient égaux :

$$\overrightarrow{MM'} = \vec{V}.$$

Si le vecteur $\vec{V}$ est nul, la translation n'est autre que la transformation identique (tout point M se transforme en lui-même).

La composée de deux translations de vecteurs $\vec{V}$ et $\vec{V'}$ est la translation de vecteur $\vec{V} + \vec{V'}$.

L'ensemble des translations, muni de la loi ∘ de composition des applications, est un groupe commutatif.

TRAPÈZE : voir Polygones

TREILLIS

Borne supérieure, borne inférieure d'un ensemble

Premier exemple

On a défini dans l'ensemble des entiers naturels positifs ou nuls $\mathbb{N}$ une relation « est plus petit que » (ou « est inférieur ou égal à ») (x est plus petit que y s'il existe un entier positif ou nul n tel que $x + n = y$). Soit X l'ensemble des nombres entiers compris entre 3 et 6 ; X est donc l'ensemble des nombres 3, 4, 5, 6 :

$$X = \{ 3, 4, 5, 6 \}.$$

X admet quatre minorants (c'est-à-dire des nombres inférieurs ou égaux à tous les nombres de X), qui sont 0, 1, 2, 3, et une infinité de majorants (c'est-à-dire des nombres supérieurs ou égaux à tous les éléments de X), qui sont tous les nombres entiers supérieurs ou égaux à 6.
Parmi tous les minorants de X, il en existe un qui est supérieur à tous les autres : c'est 3 ; 3 est la borne inférieure de X (le plus grand des minorants).
Parmi tous les majorants de X, il en existe un qui est inférieur à tous les autres : c'est 6 ; 6 est la borne supérieure de X (le plus petit des majorants).
On constate que, dans cet exemple, la borne inférieure de X ainsi que sa borne supérieure appartiennent à X.

Deuxième exemple

Soit E un ensemble, dont on appellera les parties A, B, C... ; on dira qu'une partie A de E est plus petite qu'une partie B si A est contenue (incluse) dans B. Considérons l'ensemble X constitué de deux parties A et B de E :

$$X = \{ A, B \}.$$

Parmi toutes les parties de E, il y en a qui contiennent à la fois A et B (donc qui sont plus grandes que A et B) : ce sont celles qui sont formées de tous les éléments de A, de tous les éléments de B et d'autres éléments de E ; ces parties sont des majorants de X. Parmi toutes ces parties contenant à la fois A et B, il en est une qui est la plus petite de toutes, c'est celle qui ne contient aucun autre élément en dehors de ceux de A et de B. C'est la réunion de A et de B : $A \cup B$. Cette réunion est la borne supérieure de X.
Parmi toutes les parties de E, il y en a qui sont contenues à la fois dans A et dans B (donc qui sont plus petites que A et B), ce sont celles qui sont formées d'éléments appartenant à la fois à A et à B (ces parties sont donc des minorants de X). Parmi toutes les parties contenues à la fois dans A et dans B, il y en a une qui est la plus grande de toutes : celle qui est formée de *tous* les éléments communs à A et à B, c'est-à-dire l'intersection de A et de B : $A \cap B$. Cette intersection $A \cap B$ est la borne inférieure de X.
On constate que, dans cet exemple-ci, ni la borne inférieure ni la borne supérieure de X n'appartiennent en général à X◆.

◆ Ne pas croire que, systématiquement, ou bien les deux bornes appartiennent à X, ou bien aucune des deux n'appartient à X. Il est des cas où l'une des deux appartient à X et pas l'autre.

De façon générale, soient E un ensemble muni d'une relation d'ordre et A un sous-ensemble de E :
— si l'ensemble des majorants de A admet un plus petit élément, celui-ci est appelé la borne supérieure de A (il n'appartient pas nécessairement à A mais appartient évidemment à E) ; on le note sup A ;
— si l'ensemble des minorants de A admet un plus grand élément, celui-ci est appelé la borne inférieure A (il n'appartient pas nécessairement à A mais appartient à E) ; on le note inf A.

Si on considère la partie de $\mathbb{N}$, ensemble des entiers naturels, constituée des deux éléments 3 et 6 :

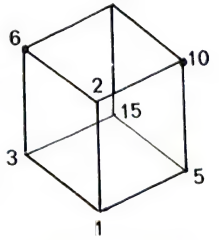
$$X = \{ 3, 6 \},$$

cette partie X admet une borne inférieure : 3, et une borne supérieure : 6 (qui sont des entiers). Il en est de même pour toute partie de $\mathbb{N}$ constituée de deux éléments x et y ; si x est inférieur à y, x est la borne inférieure et y la borne supérieure de la partie $\{ x, y \}$.
De même, si on considère deux parties A et B du même ensemble E, le nouvel ensemble X dont les éléments sont A et B admet une borne supérieure qui est $A \cup B$ et une borne inférieure qui est $A \cap B$; $A \cup B$ et $A \cap B$ n'appartiennent pas en général à X mais sont des parties de E.
$\mathbb{N}$ muni de la relation « est plus petit que » et l'ensemble $\mathfrak{P}(E)$ muni de la relation « est inclus dans » sont des *treillis*, puisqu'un treillis est un ensemble qui satisfait aux deux conditions suivantes :
— il est muni d'une relation d'ordre ;
— toute partie de cet ensemble constituée de deux éléments admet une borne inférieure et une borne supérieure qui appartiennent à l'ensemble.
La relation d'ordre choisie pour exemple sur $\mathbb{N}$ est la relation d'ordre « est plus petit que ». Il y a une autre relation d'ordre sur $\mathbb{N}^*$ ($\mathbb{N}$ sauf zéro) : la relation « divise ». Pour cette relation, $\mathbb{N}^*$ est un treillis, car, étant donné deux entiers quelconques a et b, parmi tous les nombres entiers divisés à la fois par a et par b, il en existe un qui les divise tous : c'est le P.P.C.M. de a et de b, qui est donc la borne supérieure de $X = \{ a, b \}$; parmi tous les nombres qui divisent à la fois a et b, il en est un qui est divisé par tous ces nombres : c'est le P.G.C.D. de a et de b, qui est donc la borne inférieure de X.

L'ensemble des diviseurs d'un nombre entier n constitue aussi un treillis pour la relation « divise », car, si deux nombres a et b divisent n, leur P.G.C.D. et leur P.P.C.M. divisent aussi n, donc appartiennent bien à l'ensemble considéré (voir figure ci-contre).
Si, pour la relation d'ordre R, un ensemble E est un treillis, toute partie $\{ x, y \}$ formée de deux éléments de E admet une borne inférieure notée $\inf \{ x, y \}$ (ou encore $x \wedge y$) et une borne supérieure notée $\sup \{ x, y \}$ (ou $x \vee y$). Quels que soient les éléments x, y, z de E, les propriétés suivantes sont vérifiées :

1) $x \wedge x = x$; $x \vee x = x$.
2) $x \wedge y = y \wedge x$; $x \vee y = y \vee x$.
3) $x \wedge (y \wedge z) = (x \wedge y) \wedge z$; $x \vee (y \vee z) = (x \vee y) \vee z$.
4) $x \wedge (x \vee y) = x$; $x \vee (x \wedge y) = x$.

Treillis des diviseurs de 30



Chaque trait ascendant indique que le nombre le plus bas divise le plus haut.
Par exemple, 3 divise 6.
Si un trait ascendant joint 3 à 6, un autre 6 à 30, on n'indique pas celui qui devrait joindre 3 à 30.

L'application qui à deux éléments x et y de E associe l'élément $x \wedge y$ (ou $x \vee y$) peut être considérée comme une loi de composition interne sur E puisque $x \wedge y$ (comme $x \vee y$) appartient à E. La propriété 1) traduit l'idempotence de ces lois ; la propriété 2) en traduit la commutativité et la propriété 3) l'associativité.

Treillis de Boole ◆

◆ Voir Algèbre de Boole.

Si on considère le treillis constitué par les parties d'un ensemble E :

1) Ce treillis a un élément plus grand que tous les autres : c'est l'ensemble E tout entier, et un élément plus petit que tous les autres : c'est l'ensemble vide $\varnothing$.

2) Chaque partie A de E est telle qu'il existe une partie B de E telle que la réunion de A et de B soit égale à E et que l'intersection de A et de B soit l'ensemble vide ; cette partie B n'est autre que le complémentaire de A dans E : $\complement_E A$.

$$A \cap \complement_E A = \varnothing, \qquad A \cup \complement_E A = E.$$

3) Quelles que soient les parties A, B, C de E :

— l'intersection de A avec la réunion de B et de C est égale à la réunion de l'intersection de A et de B avec l'intersection de A et de C :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

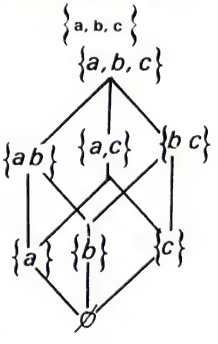
(distributivité de l'intersection par rapport à la réunion) ;

— la réunion de A avec l'intersection de B et de C est égale à l'intersection de la réunion de A et de B avec la réunion de A et de C :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

(distributivité de la réunion par rapport à l'intersection).

Treillis des parties de



Le treillis des parties de E *est un treillis de Boole.*

Plus généralement, si E est un treillis pour la relation d'ordre notée $\leqslant$, on dit que E est un treillis de Boole si les trois propriétés suivantes sont vérifiées :

1) E admet un élément u supérieur à tous les éléments de E et un élément z inférieur à tous les éléments de E.

2) Quel que soit l'élément x de E, il existe un élément x' de E ◆ tel que la borne supérieure de $\{ x, x' \}$ soit égale à u :

$$x \vee x' = u,$$

et tel que la borne inférieure de $\{ x, x' \}$ soit égale à z :

$$x \wedge x' = z.$$

◆ x' est appelé complément de x et le treillis est dit complémenté.

3) Quels que soient les éléments x, y, t de E, on a ◆ :

$x \wedge (y \vee t) = (x \wedge y) \vee (x \wedge t)$ (distributivité de $\wedge$ par rapport à $\vee$)

$x \vee (y \wedge t) = (x \vee y) \wedge (x \vee t)$ (distributivité de $\vee$ par rapport à $\wedge$).

◆ Le treillis est qualifié de distributif.

Un treillis de Boole T peut être muni de deux lois de composition interne qui en font un anneau de Boole :

— la loi notée $+$, telle que, pour tout couple d'éléments a et b de T :

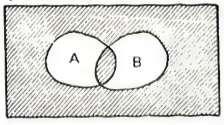
$a + b = (a' \wedge b') \vee (a \wedge b)$ (a' et b' étant les compléments de a et de b) ;

— la loi notée $\times$, telle que, pour tout couple d'éléments a et b de T :

$a \times b = a \vee b$.

L'ensemble des parties de E peut être muni de deux lois de composition interne :
— l'une, notée +, telle que : $A + B = (\complement A \cap \complement B) \cup (A \cap B)$ ◆ ;
— l'autre, notée ×, telle que : $A \times B = A \cup B$.
E, muni de ces deux lois, est un anneau de Boole.
Remarquons que l'on peut munir E de deux autres lois de composition interne : la différence symétrique Δ et l'intersection, et que, muni de ces deux lois, E est encore un anneau de Boole.

◆ Cette loi de composition est appelée somme symétrique de A et de B.



(A+B hachuré)

TRIANGLE : voir Polygones et les pages 214, 320, 326.

TRIANGLE DE PASCAL ◆ page 216.

◆ Il est absolument indispensable de lire le mot Combinaison.

Tableau constitué par les nombres $\complement_n^p$ rangés de la façon suivante : on numérote les lignes 0, 1, 2, ..., n et les colonnes 0, 1, 2, ..., p. Au croisement de la ligne n et de la colonne p, on met le nombre $\complement_n^p$.
Chaque élément de ce tableau est obtenu en faisant la somme de celui qui se trouve dans la même colonne, mais dans la ligne précédente, et de celui qui se trouve dans la ligne précédente et la colonne précédente ◆. Il n'y a que des 1 dans la 1re colonne $\left(\complement_n^0 = 1\right)$, que des 1 sur la diagonale $\left(\complement_n^n = 1\right)$ et il n'y a des nombres qu'en dessous de la diagonale, car, pour chaque valeur de n, p ne peut prendre que des valeurs comprises entre 0 et n.

◆ Cela grâce à la formule $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$ énoncée au mot Combinaison.

n \ p	0	1	2	3	4	5	6	7	8	9	0
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

TRIANGULAIRE (Matrice)

Une matrice triangulaire supérieure est une matrice carrée dont tous les termes situés en dessous de la diagonale principale sont nuls ; elle est triangulaire inférieure si ce sont tous les termes situés au-dessus de la diagonale principale qui sont nuls.

Matrice triangulaire supérieure

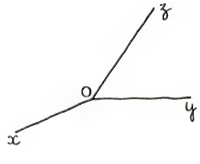
$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 5 \\ 0 & 0 & 2 \end{pmatrix}$$

Matrice triangulaire inférieure

$$\begin{pmatrix} -1 & 0 & 0 \\ 5 & 2 & 0 \\ 3 & 6 & 0 \end{pmatrix}$$

TRIÈDRE

Figure formée par trois demi-droites, non coplanaires, de même origine. Les trois plans xOy, yOz et zOx sont appelés les faces du trièdre.
Si Ox, Oy, Oz sont deux à deux perpendiculaires, le trièdre est trirectangle.



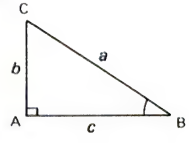
TRIGONOMÉTRIE pages 159, 289, 297, 298, 326, 328, 453.

La trigonométrie est la branche des mathématiques qui étudie les fonctions trigonométriques, les relations entre ces fonctions, les relations entre les côtés et les angles d'un triangle et leurs applications à différents problèmes. La trigonométrie a de nombreuses applications en physique, en navigation, en construction, etc.

Relations entre les côtés et les angles d'un triangle

Relations trigonométriques dans un triangle rectangle

En trigonométrie élémentaire, on définit les sinus et les cosinus des angles aigus d'un triangle rectangle de la façon suivante : le cosinus est égal au rapport de la longueur du côté adjacent à la longueur de l'hypoténuse et le sinus est égal au rapport de la longueur du côté opposé à la longueur de l'hypoténuse. Ainsi, dans l'exemple ci-contre :



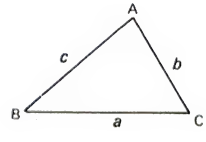
$$\cos B = \frac{c}{a} \qquad \sin B = \frac{b}{a}.$$

On appelle tangente d'un angle aigu le rapport de la longueur du côté opposé à la longueur du côté adjacent :

$$\text{tg } B = \frac{b}{c}, \text{ donc } \text{tg } B = \frac{\sin B}{\cos B}.$$

Relations trigonométriques dans un triangle quelconque

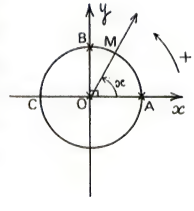
Dans la figure ci-contre :



$$a^2 = b^2 + c^2 - 2\,bc \cos A$$
$$b^2 = a^2 + c^2 - 2\,ac \cos B$$
$$c^2 = a^2 + b^2 - 2\,ab \cos C$$
$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}.$$

Fonctions trigonométriques

Soient un cercle de centre O et de rayon 1, un axe Ox (un axe Oy perpendiculaire) et un sens de parcours du cercle (sens inverse de celui du déplacement des aiguilles d'une montre) choisi pour sens positif. Ce cercle est appelé cercle trigonométrique. A chaque nombre réel x on fait correspondre un point M du cercle trigonométrique (le point tel que l'angle orienté de Ox avec $\overrightarrow{OM}$ ait une mesure en radians égale à x).

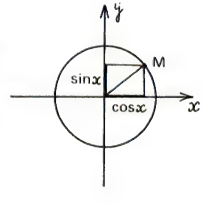


Ainsi, au nombre réel 0 correspond le point A, à $\frac{\pi}{2}$ correspond B, à π correspond C, à 2π correspond A, etc.

Par définition, on appelle :

— *cosinus du nombre réel x* la mesure algébrique de la projection sur Ox du vecteur $\overrightarrow{OM}$;

— *sinus du nombre réel x* la mesure algébrique de la projection sur Oy du vecteur $\overrightarrow{OM}$.



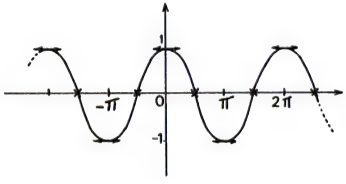
A partir du sinus et du cosinus de x on peut définir la tangente du nombre x comme étant le rapport de sin x à cos x : $\text{tg } x = \frac{\sin x}{\cos x}$.

Ces fonctions sinus et cosinus sont définies pour tout x réel ; elles sont continues, périodiques de période 2π ; les valeurs qu'elles prennent sont comprises entre -1 et $+1$; leurs dérivées sont :

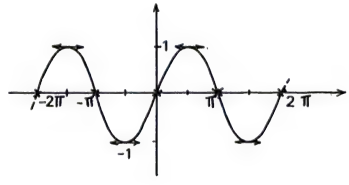
$$(\cos x)' = -\sin x \quad \text{et} \quad (\sin x)' = \cos x.$$

Les courbes représentatives de ces fonctions sont des sinusoïdes :

Courbe représentative de la fonction cosinus



Courbe représentative de la fonction sinus



Relations entre les différentes fonctions trigonométriques

$$\sin^2 x + \cos^2 x = 1 \qquad 1 + \operatorname{tg}^2 x = \frac{1}{\cos^2 x} \qquad 1 + \operatorname{cotg}^2 x = \frac{1}{\sin^2 x}.$$

Formules d'addition

$$\cos(a+b) = \cos a \cos b - \sin a \sin b \qquad \cos(a-b) = \cos a \cos b + \sin a \sin b$$
$$\sin(a+b) = \sin a \cos b + \sin b \cos a \qquad \sin(a-b) = \sin a \cos b - \sin b \cos a$$

$$\operatorname{tg}(a+b) = \frac{\operatorname{tg} a + \operatorname{tg} b}{1 - \operatorname{tg} a \operatorname{tg} b} \qquad \operatorname{tg}(a-b) = \frac{\operatorname{tg} a - \operatorname{tg} b}{1 + \operatorname{tg} a \operatorname{tg} b}$$

$$\begin{aligned} \cos 2a &= \cos^2 a - \sin^2 a \\ &= 2\cos^2 a - 1 \\ &= 1 - 2\sin^2 a \end{aligned}$$

$$\sin 2a = 2 \sin a \cos a$$

$$\operatorname{tg} 2a = \frac{2 \operatorname{tg} a}{1 - \operatorname{tg}^2 a}$$

$$\cos a = \frac{1 - t^2}{1 + t^2}, \quad \sin a = \frac{2t}{1 + t^2}, \quad \operatorname{tg} a = \frac{2t}{1 - t^2} \quad \text{si on a posé } t = \operatorname{tg} \frac{a}{2}$$

Formules de transformation de sommes en produits

$$\cos a + \cos b = 2 \cos \frac{a+b}{2} \cos \frac{a-b}{2} \qquad \sin a + \sin b = 2 \sin \frac{a+b}{2} \cos \frac{a-b}{2}$$

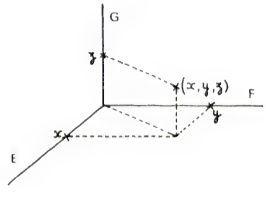
$$\cos a - \cos b = -2 \sin \frac{a+b}{2} \sin \frac{a-b}{2} \qquad \sin a - \sin b = 2 \sin \frac{a-b}{2} \cos \frac{a+b}{2}$$

$$\operatorname{tg} a + \operatorname{tg} b = \frac{\sin(a+b)}{\cos a \cos b} \qquad \operatorname{tg} a - \operatorname{tg} b = \frac{\sin(a-b)}{\cos a \cos b}$$

◆ Lire d'abord Cartésien et Couple.

TRIPLET ◆ pages 60, 61, 367.

Soient E, F, G trois ensembles ; on appelle triplet tout élément (x, y, z) du produit cartésien $E \times F \times G$ ($x \in E$, $y \in F$, $z \in G$).



TRIRECTANGLE : voir Trièdre

TRIVIAL : synonyme de banal, évident.

U à Z

UNIFORME, UNIFORMÉMENT

Continuité uniforme

Une fonction réelle f d'une variable réelle x est continue en un point x_0 de son domaine de définition si à tout nombre réel positif ε (aussi petit que l'on veut) on peut faire correspondre un nombre réel positif α tel que, si la différence $x - x_0$ est, en valeur absolue, inférieure à α, alors la différence $f(x) - f(x_0)$ est, en valeur absolue, inférieure à ε :

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ |x - x_0| < \alpha \Longrightarrow |f(x) - f(x_0)| < \varepsilon.$$

ε et x_0 étant donnés, le nombre α dépend à priori de ces deux nombres ε et x_0. On dit que la fonction est uniformément continue en x_0 si le nombre α ne dépend pas de x_0 mais ne dépend que de ε.

Ainsi, la fonction f telle que $f(x) = 2x + 3$ est ◆ continue en n'importe quelle valeur x_0 de la variable, car :

◆ Définie pour toute valeur réelle de x.

$$|x - x_0| < \frac{\varepsilon}{2} \Longrightarrow |f(x) - f(x_0)| = |(2x + 3) - (2x_0 + 3)|$$

$$= |2(x - x_0)| = 2|x - x_0| < \varepsilon.$$

Le nombre $\alpha = \dfrac{\varepsilon}{2}$ ne dépend pas de x_0, donc la continuité est uniforme sur tout $\mathbb{R}$ (ensemble de nombres réels).

Par contre, la fonction g telle que $g(x) = x^2$ est continue pour n'importe quelle valeur de x_0 mais n'est pas uniformément continue.

Toute fonction uniformément continue est continue, mais la réciproque est fausse en général (voir l'exemple de $g(x) = x^2$, qui est continue mais non uniformément continue). La réciproque n'est vraie que si la fonction est définie sur un intervalle fermé borné (c'est le théorème de Heine).

Convergence uniforme

Une suite dont le terme général est une fonction de x : $u_n(x)$, converge vers une limite $u(x)$ quand n tend vers l'infini si, quel que soit le nombre ε donné, positif et aussi petit que l'on veut, on peut rendre la différence $u_n(x) - u(x)$, en valeur absolue, inférieure à ε, à condition de choisir le nombre n supérieur à un certain entier N. Cet entier N dépend de x et de ε ; s'il ne dépend que de ε, on dit que la convergence est uniforme sur l'ensemble de définition de tous les $u_n(x)$.

La suite définie par $u_n(x) = \dfrac{x}{n}$, où x est un nombre réel strictement positif,

converge et a pour limite 0 quand n tend vers $+\infty$, mais ne converge pas uniformément, car, pour rendre $\left|\frac{x}{n} - 0\right|$ inférieur à ε, c'est-à-dire $\frac{x}{n} < \varepsilon$, il faut que n soit supérieur à $\frac{x}{\varepsilon}$. Ce nombre $\frac{x}{\varepsilon}$ dépend de x ; la convergence n'est pas uniforme.

UNITAIRE (Anneau) : voir Anneau, l'article « l'Algèbre », pages 47 et 48, et la page 212.

VALEUR ABSOLUE

D'un nombre réel

On appelle valeur absolue d'un nombre réel x le nombre réel positif, noté $|x|$, qui est égal à :

$$\begin{cases} x, \text{ si } x \text{ est positif ou nul} \blacklozenge \\ -x, \text{ si } x \text{ est négatif} \end{cases}$$

◆ La valeur absolue de 0 est 0 : $|0| = 0$.

Ainsi, la valeur absolue de $+2$ est $+2$, celle de -2 est $-(-2)$, c'est-à-dire $+2$ également (deux nombres opposés ont même valeur absolue).

Si x et y sont deux nombres réels, la valeur absolue du produit $x.y$ est égale au produit des valeurs absolues de x et de y ; par contre, la valeur absolue de la somme $x + y$ n'est pas égale à la somme des valeurs absolues de x et de y mais lui est inférieure ou égale (il n'y a égalité que dans le cas où x et y sont de même signe) ◆.

◆ $|x \times y| = |x| \times |y|$.
$|x + y| \leqslant |x| + |y|$.

Exemples :

$$\begin{cases} x = +2, \quad y = -3 \quad x + y = -1 \quad \text{donc} \quad |x + y| = 1 \\ |x| = 2, \quad |y| = 3 \qquad\qquad\qquad\qquad\quad |x| + |y| = 5 \end{cases}$$

$$\begin{cases} x = -2, \quad y = -3 \quad x + y = -5 \quad \text{donc} \quad |x + y| = 5 \\ |x| = 2, \quad |y| = 3 \qquad\qquad\qquad\qquad\quad |x| + |y| = 5 \end{cases}$$

En algèbre : sur un corps

Une valeur absolue est une application (notée $x \longmapsto |x|$) de ce corps dans $\mathbb{R}^+$ ◆ qui satisfait aux conditions :

◆ $\mathbb{R}^+$: ensemble des nombres réels positifs ou nuls. Un corps est muni d'une addition et d'une multiplication notées $+$ et $\times$; 0 est l'élément neutre de l'addition.

1) La valeur absolue est nulle ◆ si et seulement si x est nul.

◆ Nul veut dire égal à 0.

2) Pour tout couple (x, y) d'éléments de K, la valeur absolue du produit $x \times y$ est égale au produit des valeurs absolues ◆.

◆ $|x \times y| = |x| \times |y|$.

3) Pour tout couple (x, y), la valeur absolue de la somme $x + y$ est inférieure ou égale à la somme des valeurs absolues ◆. Cette dernière condition est connue sous le nom d'inégalité triangulaire.

◆ $|x + y| \leqslant |x| + |y|$.

Par exemple, sur $\mathbb{C}$ (le corps des nombres complexes), on peut définir une valeur absolue : l'application qui à tout nombre complexe $x + iy$ (x et y réels) fait correspondre le nombre réel positif $\sqrt{x^2 + y^2}$ satisfait aux trois conditions, donc est une valeur absolue sur $\mathbb{C}$.

Valeur absolue triviale :

Pour tout $x \neq 0 \quad |x| = 1$

$x = 0 \quad |x| = 0$

C'est bien une valeur absolue (elle vérifie les trois conditions).

VALEUR PROPRE, VECTEUR PROPRE (d'une matrice)

Étant donné une matrice carrée d'ordre n à coefficients réels (ou complexes) M, un vecteur $\vec{V} = (x_1, x_2, \ldots, x_n)$, non nul, est un vecteur propre de la matrice s'il existe un nombre réel (ou complexe) λ tel que le produit de la matrice M par la matrice $V = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$ soit égal au produit du nombre λ par la matrice V :

$$M \times V = \lambda \times V.$$

Le nombre λ est appelé valeur propre de M, associée au vecteur propre $\vec{V}$. Ainsi, la matrice $M = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}$ admet pour vecteurs propres :

$\vec{V_1}$ de composantes (1, — 1), car $\begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \end{pmatrix} = 4 \times \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$$M \times V_1 = 4 \times V_1$$

$\vec{V_2}$ de composantes (3, 2), car $\begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \end{pmatrix} = (-1) \times \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

$$M \times V_2 = (-1) \times V_2.$$

Les valeurs propres associées sont respectivement 4 et — 1.
Tous les vecteurs colinéaires à V_1 sont aussi vecteurs propres de M, ainsi que tous les vecteurs colinéaires à V_2.

Pour trouver $\vec{V_1}$ et $\vec{V_2}$, on a dû chercher x et y (non tous deux nuls) tels que :

$$\begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix};$$

autrement dit, on a dû résoudre le système (S) :

$$\begin{cases} 1x - 3y = \lambda x \\ -2x + 2y = \lambda y \end{cases} \quad \text{soit} \quad \begin{cases} (1 - \lambda)x - 3y = 0 \\ -2x + (2 - \lambda)y = 0 \end{cases}$$

Ce système ne devant pas admettre la solution $x = 0$, $y = 0$, son déterminant $\begin{vmatrix} 1-\lambda & -3 \\ -2 & 2-\lambda \end{vmatrix}$ doit être nul ◆, donc : $(1 - \lambda)(2 - \lambda) - (-3)(-2) = 0$,

◆ Cette équation s'appelle équation caractéristique de M.

ce qui nous permet de trouver $\lambda = 4$ et $\lambda = -1$. On a ensuite remplacé successivement λ par 4 et par — 1 dans le système (S) :

$$\lambda = 4 : \begin{cases} -3x - 3y = 0 \\ -2x - 2y = 0 \end{cases} \iff \begin{cases} -3(x + y) = 0 \\ -2(x + y) = 0. \end{cases}$$

La seconde équation est proportionnelle à la première ; x et y doivent être opposés. Tous les vecteurs propres correspondant à $\lambda = 4$ ont des composantes de la forme $(x, -x)$.

$$\lambda = -1 : \quad \begin{matrix} 2x - 3y = 0 \\ -2x + 3y = 0. \end{matrix} \qquad x \text{ et } y \text{ doivent être tels que } 2x = 3y.$$

VARIABLE (aléatoire) : voir l'article « les Probabilités », pages 424 et 425.

VARIABLE RÉDUITE : voir l'article « les Statistiques », pages 470, 471.

VARIANCE : voir l'article « les Statistiques », pages 468 à 471, 485, 511 à 514 et les pages 235, 430, 434.

VECTEUR : voir l'article « l'Algèbre linéaire », page 55 et les pages 183, 329.

VECTORIEL (Espace) : voir l'article « l'Algèbre linéaire », pages 55 à 76 et les pages 211, 314, 330.

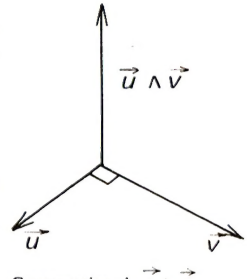
VECTORIEL (Produit) pages 390, 391, 396, 410, 443, 450, 520.

Si $\vec{u}$ et $\vec{v}$ sont deux vecteurs de l'espace de composantes respectives (x, y, z) et (x', y', z'), le produit vectoriel de ces deux vecteurs est le vecteur dont les composantes sont :

$$X = yz' - y'z\,;\ Y = zx' - z'x\,;\ Z = xy' - x'y.$$

On le note $\vec{u} \wedge \vec{v}$.

Si l'un des deux vecteurs $\vec{u}$ et $\vec{v}$ est nul, ou colinéaire à l'autre, le produit vectoriel $\vec{u} \wedge \vec{v}$ est nul. Sinon, c'est un vecteur perpendiculaire au plan des deux vecteurs $\vec{u}$ et $\vec{v}$, tel que le trièdre $\left(\vec{u}, \vec{v}, \vec{u} \wedge \vec{v}\right)$ soit de sens direct ; sa longueur est égale au produit des longueurs de $\vec{u}$ et de $\vec{v}$ par la valeur absolue du sinus de l'angle des deux vecteurs $\vec{u}$ et $\vec{v}$.



Construction de $\vec{u} \wedge \vec{v}$, dans le cas où $\vec{u}$ et $\vec{v}$ sont perpendiculaires et de longueurs respectives 1 et 2.

1) $\vec{u} \wedge \vec{v} = -\vec{v} \wedge \vec{u}$.

2) Ce produit est distributif par rapport à la somme :

$$\vec{u} \wedge \left(\vec{v} + \vec{w}\right) = \vec{u} \wedge \vec{v} + \vec{u} \wedge \vec{w}.$$

VENN (Diagramme de) : voir Diagramme et la page 275.

VIDE pages 263, 300, 336.

L'ensemble vide est un ensemble qui n'a aucun élément ; cet ensemble est unique et on le note $\emptyset$.

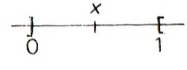
VOISINAGE ◆ pages 330, 345, 531.

◆ Vocabulaire de topologie. Voir Topologie et Intervalle.

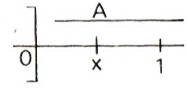
Un voisinage d'un point x d'un espace topologique E est une partie de E contenant un ouvert auquel x appartient.

Exemples

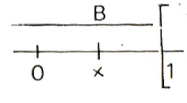
Si x est un nombre réel de l'intervalle ouvert $]0, 1[$, l'ensemble des nombres y supérieurs ou égaux à 0 et strictement inférieurs à 1 est un voisinage de x (puisqu'il contient l'intervalle ouvert $]0, 1[$ auquel x appartient).



— L'ensemble A des nombres y supérieurs ou égaux à 0 est un voisinage de x (puisqu'il contient l'ouvert $]0, 1[$ qui contient x).



— L'ensemble B des nombres y inférieurs ou égaux à 1 est un voisinage de x (puisqu'il contient $]0, 1[$ qui contient x).



— Par contre, l'ensemble C de tous les nombres y qui sont différents de x n'est

pas un voisinage x (il ne peut contenir un ouvert contenant x puisque x ne lui appartient pas).
— L'ensemble D des nombres y supérieurs ou égaux à x n'est pas un voisinage de x (il ne contient pas d'intervalle ouvert contenant x, car un intervalle ouvert contenant x contient aussi des nombres inférieurs à x et ces nombres inférieurs à x ne font pas partie de D).

Propriétés

Un sous-ensemble d'un espace topologique E est un ouvert de E si et seulement s'il est un voisinage de chacun de ses points.
Tout voisinage de x contient x.
Toute partie de E contenant un voisinage de x est un voisinage de x.
Toute intersection d'un nombre fini de voisinages de x est un voisinage de x.
Tout voisinage V de x contient un voisinage U de x tel que V soit voisinage de chacun des points de U.

Système fondamental de voisinages

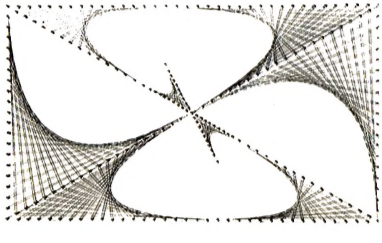
Si $\mathfrak{V}(x)$ est la famille des voisinages du point x de l'espace topologique E, on appelle système fondamental de voisinage de x toute partie $\mathfrak{U}(x)$ de $\mathfrak{V}(x)$ telle que tout élément V de $\mathfrak{V}(x)$ contienne un élément U de $\mathfrak{U}(x)$.

Exemple : les intervalles ouverts $\left]x - \frac{1}{n}, x + \frac{1}{n}\right[$, où n est un entier positif, non nul, constituent un système fondamental de voisinages du point x de la droite numérique $\mathbb{R}$, car, quel que soit le voisinage V de x, il existe un intervalle de la forme $\left]x - \frac{1}{n}, x + \frac{1}{n}\right[$ qui soit contenu dans ce voisinage.

VOLUME page 317.

Le volume d'un cube, dont la longueur de chaque côté est a, est, par définition, a^3.
Le volume d'un ensemble E de points limité par une surface S est la borne inférieure de la somme des volumes des cubes (disjoints deux à deux) qui, ensemble, contiennent tous les points de E.
On peut calculer le volume de quelques ensembles limités par des surfaces simples à l'aide du calcul intégral.

ZÉRO : voir l'article « les Ensembles », page 263 et la page 326.



Ensembles de nombres

$\mathbb{N}$ entiers naturels ; $\mathbb{Z}$ entiers relatifs ; $\mathbb{Q}$ rationnels ; $\mathbb{R}$ réels ; $\mathbb{C}$ complexes

Logique

$\neg$ négation ; et *ou* $\wedge$ conjonction ; ou *ou* $\vee$ disjonction

$\Longrightarrow$ implique ; $\Longleftrightarrow$ équivalent à

$\forall$ quel que soit ; $\exists$ il existe au moins un

Ensembles

$\varnothing$ ensemble vide ; $\in$ appartient à ; $\notin$ n'appartient pas à

$\subset$ est inclus dans ; $\not\subset$ n'est pas inclus dans ; $\cup$ réunion ; $\cap$ intersection

E^n produit cartésien de n ensembles égaux à E

$E \Delta F$ différence symétrique des deux ensembles E et F

$\complement_E F$ complémentaire de F dans E

Relations

$x \equiv y$ (mod. $\mathcal{R}$) x congru à y modulo $\mathcal{R}$

$E/\mathcal{R}$ ensemble quotient de l'ensemble E par la relation d'équivalence $\mathcal{R}$

$\leqslant$ inférieur ou égal à ; $<$ strictement inférieur à

$\geqslant$ supérieur ou égal à ; $>$ strictement supérieur à

$x \mathcal{R} y$ x est lié à y par la relation $\mathcal{R}$

$f : x \longmapsto f(x)$ x a pour image $f(x)$ par l'application f

Lois de composition

$\ast$, T, $\perp$, $\circ$ étoile, truc, antitruc, rond

Autres symboles

$\sum_{i=1}^{n} x = x_1 + x_2 + \ldots + x_n$ somme des n éléments $x_1, x_2, \ldots, x_n$

$\prod_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n$ produit des n éléments $x_1, x_2, \ldots, x_n$

$\frac{df}{dx}$ *ou* $f'(x)$ dérivée de f ; f'' dérivée seconde de f ; $f^{(n)}$ dérivée n-ième de f

df différentielle de f ; df_{x_0} *ou* $d_{x_0} f$ différentielle de f au point x_0

$\int f(x)\,dx$ intégrale de f ; $\iint$ intégrale double ; $\iiint$ intégrale triple

$|x|$ valeur absolue de x ou module de x ; $\|x\|$ norme de x

$n!$ factorielle n (lire le mot au dictionnaire)

e élément neutre, base de l'exponentielle

f, g, h application, fonction

f^{-1} application réciproque de l'application f

$f(x)$ image de l'élément x par l'application f

I ensemble d'indices

$\mathcal{F}$ filtre

$\alpha, \beta, \lambda, \mu$ scalaires

σ permutation, écart type

Cet ouvrage a été imprimé sur les presses des Petits-Fils de Léonard Danel, maîtres imprimeurs à Loos-lez-Lille
N° d'éditeur 81, n° d'imprimeur 8544.74.6
La structure Algo-livre, décrite pages 6 et 7, est protégée par plusieurs brevets.